FaceAge, a deep learning system to estimate biological age from face photographs to improve prognostication: a model development and validation study

Dennis Bontempi, Osbert Zalay*, Danielle S Bitterman, Nicolai Birkbak, Derek Shyr, Fridolin Haugg, Jack M Qian, Hannah Roberts, Subha Perni, Vasco Prudente, Suraj Pai, Andre Dekker, Benjamin Haibe-Kains, Christian Guthier, Tracy Balboni, Laura Warren, Monica Krishan, Benjamin H Kann, Charles Swanton, Dirk De Ruysscher, Raymond H Mak†, Hugo J W L Aerts†*

Summary

Background As humans age at different rates, physical appearance can yield insights into biological age and physiological health more reliably than chronological age. In medicine, however, appearance is incorporated into medical judgements in a subjective and non-standardised way. In this study, we aimed to develop and validate FaceAge, a deep learning system to estimate biological age from easily obtainable and low-cost face photographs.

Methods FaceAge was trained on data from 58 851 presumed healthy individuals aged 60 years or older: 56 304 individuals from the IMDb–Wiki dataset (training) and 2547 from the UTKFace dataset (initial validation). Clinical utility was evaluated on data from 6196 patients with cancer diagnoses from two institutions in the Netherlands and the USA: the MAASTRO, Harvard Thoracic, and Harvard Palliative cohorts FaceAge estimates in these cancer cohorts were compared with a non-cancerous reference cohort of 535 individuals. To assess the prognostic relevance of FaceAge, we performed Kaplan–Meier survival analysis and Cox modelling, adjusting for several clinical covariates. We also assessed the performance of FaceAge in patients with metastatic cancer receiving palliative treatment at the end of life by incorporating FaceAge into clinical prediction models. To evaluate whether FaceAge has the potential to be a biomarker for molecular ageing, we performed a gene-based analysis to assess its association with senescence genes.

Findings FaceAge showed significant independent prognostic performance in various cancer types and stages. Looking older was correlated with worse overall survival (after adjusting for covariates per-decade hazard ratio [HR] 1·151, p=0·013 in a pan-cancer cohort of n=4906; 1·148, p=0·011 in a thoracic cohort of n=573; and 1·117, p=0·021 in a palliative cohort of n=717). We found that, on average, patients with cancer looked older than their chronological age (mean increase of 4·79 years with respect to non-cancerous reference cohort, p<0·0001). We found that FaceAge can improve physicians' survival predictions in patients with incurable cancer receiving palliative treatments (from area under the curve 0·74 [95% CI 0·70–0·78] to 0·8 [0·76–0·83]; p<0·0001), highlighting the clinical use of the algorithm to support end-of-life decision making. FaceAge was also significantly associated with molecular mechanisms of senescence through gene analysis, whereas age was not.

Interpretation Our results suggest that a deep learning model can estimate biological age from face photographs and thereby enhance survival prediction in patients with cancer. Further research, including validation in larger cohorts, is needed to verify these findings in patients with cancer and to establish whether the findings extend to patients with other diseases. Subject to further testing and validation, approaches such as FaceAge could be used to translate a patient's visual appearance into objective, quantitative, and clinically valuable measures.

Funding US National Institutes of Health and EU European Research Council.

Introduction

Emerging evidence suggests that people age at different rates. Interpersonal differences in genetic and lifestyle factors such as diet, stress, smoking, and alcohol use have been shown to influence the ageing process and affect DNA methylation status,[1–3] telomere length,[4–6] and gene and protein expression patterns.[7–9] There is no single clock that measures biological age directly, but establishing biomarkers that correlate with survival time (ie, time until death) could have clinically relevant applications. Finding an appropriate surrogate of a person's biological age could provide a better predictor of their physiological health and life expectancy than chronological age. This is especially important in medicine, in which both diseases and treatments can cause cellular damage and accelerate the ageing process, and an accurate estimation of biological age could support treatment decisions and allow better quantification of the relative risk–benefit ratio of proposed treatments. For

Lancet Digit Health 2025

Published Online https://doi.org/10.1016/j.landig.2025.03.002

*Joint first authors
†Joint last authors

Artificial Intelligence in Medicine Program, Mass General Brigham, Harvard Medical School, Boston, MA, USA (D Bontempi PhD, O Zalay PhD, D S Bitterman MD, F Haugg MSc, J M Qian MD, H Roberts MD, S Perni MD, V Prudente MSc, S Pai MSc, C Guthier PhD, B H Kann MD, R H Mak MD, Prof H J W L Aerts PhD); **Department of Radiation Oncology, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA** (D Bontempi, O Zalay, D S Bitterman, F Haugg, J M Qian, H Roberts, S Perni, V Prudente, S Pai, C Guthier, T Balboni MD, L Warren MD, M Krishan MD, B H Kann, R H Mak, Prof H J W L Aerts); **Department of Radiology and Nuclear Medicine, CARIM & GROW, Maastricht University, Maastricht, Netherlands** (D Bontempi, V Prudente, S Pai, Prof H J W L Aerts); **Department of Radiation Oncology (MAASTRO), Maastricht University, Maastricht, Netherlands** (D Bontempi, Prof A Dekker PhD, Prof D De Ruysscher MD); **Division of Radiation Oncology, Queen's University, Kingston, ON, Canada** (O Zalay); **Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark** (N Birkbak PhD); **Department of Clinical Medicine and Bioinformatics Research Center, Aarhus University, Aarhus, Denmark** (N Birkbak); **Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA, USA** (D Shyr PhD); **Princess Margaret**

Cancer Centre, University Health Network, Toronto, ON, Canada (B Haibe-Kains PhD); **Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada** (B Haibe-Kains); **Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK** (Prof C Swanton MD); **Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK** (Prof C Swanton); **Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA** (Prof H J W L Aerts)

Correspondence to: Prof Hugo J W L Aerts, Artificial Intelligence in Medicine Program, Mass General Brigham, Harvard Medicine School, Boston, MA 02115, USA
haerts@bwh.harvard.edu

Research in context

Evidence before this study

On July 1, 2024, we searched PubMed using the following search terms: "(face pictures OR facial image) AND (artificial intelligence OR deep learning OR analysis) AND (aging OR cancer)". We focused on articles published from Jan 1, 2014, to Dec 20, 2023. No language restrictions were applied. Our search, which identified 1729 articles, yielded several studies in which artificial intelligence was used to estimate apparent age from face pictures, but no relevant studies in which the application of this technology to predict clinical outcomes was investigated.

Added value of this study

To our knowledge, this is the first study to validate a deep learning pipeline for the estimation of biological age from face pictures (FaceAge) and to explore the association of estimated age with clinical outcomes. We assessed the prognostic relevance of FaceAge through survival analysis and prediction modelling in various cancer types and stages, taking several confounders into account. As a proof of concept, we tested the pipeline in a clinical context by incorporating the predictions into a validated clinical model and running a survey among physicians to assess its use in clinical decision making. Finally, we performed a gene-based analysis to measure the association between FaceAge and senescence genes to evaluate whether it could potentially be a biomarker for molecular ageing.

Implications of all the available evidence

Our study builds on previous investigations of facial age as a potential marker of biological age. Our results suggest that the facial characteristics visible in a photograph hold information about a person's age that deep learning algorithms can use to enhance the accuracy of survival forecasts for patients with cancer. On average, the facial age of patients with cancer was approximately 5 years older than their chronological age. Moreover, patients with cancer had significantly older facial age than healthy controls and patients without cancer treated for conditions that are benign or precancerous. Further research is needed to explore the possibility of using the pipeline, which relies on easily obtainable face photographs, to improve on the current standard of subjective visual assessment routinely used in the clinic.

example, a fit 75-year-old whose biological age is 10 years younger than their chronological age might tolerate and respond to treatment better and live longer than a 60-year-old whose biological age is 10 years older than their chronological age.

In current clinical practice, a physician's overall impression of a patient constitutes an integral part of the physical examination and has a major role in clinical decision making—in estimating prognosis and in weighing the benefits and risks of diagnostic procedures and treatment. However, this is an exceptionally subjective assessment of functional status or frailty and only a rough estimation of the biological age of a patient.[10] Especially in oncology, in which the therapeutic window is often narrow and the treatment itself can worsen mortality rates, a decision to treat requires accurate estimates of whether the patient would be healthy enough to tolerate treatment and live long enough to benefit from it. Moreover, we might expect a greater biological age for patients with cancer because of the combined effect of the disease and treatment toxicity. Unfortunately, oncologists have to make these complex treatment decisions without knowing the exact biological age of a patient, relying instead on subjective performance status estimates, which contribute to a well documented poor ability to predict outcomes for their patients.[11–13] Therefore, there is a compelling need for quantitative methods to improve patient stratification and support physicians in this complex decision-making process for appropriate treatment selection. An objective measure of biological age could also allow more accurate, objective stratification within trials and better translation of results to patients in the real world. Furthermore, this approach could help to decipher the biological processes associated with premature ageing and identify individuals who age faster and are at increased risk of diseases.

In the past few years, the application of deep learning for age estimation from face images has gained considerable attention in the academic community, with several publications presenting models capable of accurately predicting an individual's age based on face photographs.[14,15] With applications mainly outside of medicine, researchers have explored various aspects of deep learning to enhance the accuracy and applicability of age estimation models.[16] For instance, the work of Rothe and colleagues[17] not only contributed to methodological advancements, but also shared important data resources publicly. To the best of our knowledge, no previous work has applied this approach in a clinical context or investigated the prognostic value of age estimation by deep learning models.

We hypothesised that a person's biological age is reflected in their facial characteristics and that deep learning algorithms can capture this information automatically from easily obtainable photographs. Such an approach could provide a more precise measure of a patient's physiological status than chronological age, providing essential information for precision medicine as an actionable clinical biomarker and prognostication factor. Early evidence for this was established in a study by Xia and colleagues,[18] in which perceived age was estimated from the faces of healthy individuals using a specialised three-dimensional imaging device and the data were shown to be associated with molecular markers of ageing. In the current study, we aimed to develop a deep learning system to estimate a person's biological age from easily obtainable face photographs, and

to assess the clinical value of these age estimates in predicting survival outcomes in patients with cancer diagnoses.

Methods

Datasets

A detailed description of the datasets used in this study can be found in the appendix (pp 2–4). We trained our deep learning pipeline, FaceAge, on a curated subset of the IMDb–Wiki database,[17] a large and publicly available age-labelled face pictures database. A total of 56 304 images from the database were selected after applying exclusion criteria, using randomisation and augmentation with rebalancing. Although ground truth is available for all individuals, and since manual curation is a cumbersome process, we optimised our experimental design for the age range typical for clinical oncology populations and inspected and assessed the quality of face photographs only for individuals aged 60 years or older (appendix pp 21–22). Patients younger than 60 years represent the minority in the clinical cohorts (appendix pp 2, 24). After a first validation on a curated subset of the UTK dataset[19] (appendix pp 2, 23), we validated the FaceAge pipeline on four different clinical cohorts: the MAASTRO cohort, a dataset of patients with a cancer diagnosis collected at the MAASTRO radiation oncology clinic, Maastricht, the Netherlands (appendix pp 3, 24); the Harvard Thoracic cohort, consisting of patients with thoracic cancer who received radiotherapy treatment at the Dana-Farber Cancer Institute and Brigham and Women's Hospital in Boston, MA, USA (appendix pp 3, 25); the Harvard Palliative cohort of patients with metastatic disease, seen for consideration of palliative-intent treatment at the Dana-Farber Cancer Institute and Brigham and Women's Hospital (appendix pp 3, 26); and the Harvard non-cancerous cohort, including patients diagnosed with benign tumour and patients diagnosed with ductal carcinoma in situ (appendix p 4). For all the clinical validation datasets, the face pictures were taken by clinical staff members at the time of a medical appointment using different commercially available digital cameras in a semi-standardised way—ie, no professional lighting was used and the background, image resolution, and patient's expression changed slightly from time to time (appendix pp 3–4, 6).

This study adheres to the ethical principles for human research outlined in the Declaration of Helsinki, and the study and its protocols were approved by the independent hospital ethics review boards (IRB) at Mass General Brigham and the Dana-Farber Harvard Cancer Center, Boston, MA, USA, and Maastricht University, Maastricht, the Netherlands. All clinical data were handled in compliance with respective institutional research policies. Specifically, the patients' data included in the MAASTRO Biobank were prospectively collected, and consent was given for the data collection. For the Harvard datasets, the data were collected retrospectively, and the IRB approved a waiver of consent (Mass General Brigham IRB protocol 2020P002969 and Dana-Farber Harvard Cancer Center IRB protocols 13-055 and 17-669). The training cohorts (IMDb–Wiki and UTKFace) are public datasets that are made available for non-commercial academic research purposes. No patient data were used to train the algorithm.

FaceAge deep learning pipeline

The FaceAge deep learning pipeline involves two main stages: face detection and feature extraction. The first stage uses a cascaded convolutional neural network[20] to locate and preprocess the face, achieving a test accuracy of 95%. The second stage uses an Inception-ResNet v1 convolutional neural network[21] to encode the face into a feature vector and perform age prediction through regression. The model performance was good for the clinically relevant age range (ie, 60 years and older) that underwent manual curation and quality assurance (mean absolute error 4·09 years), on par or better when compared with state-of-the-art models fine-tuned on the same data (appendix pp 30–34). More detailed information on the models' architecture and pipeline development can be found in the appendix (pp 18, 27–28).

See Online for appendix

Statistical and survival analysis

The pipeline's age estimation performance was validated on the UTKFace dataset[19] and on the aforementioned clinical datasets by comparing FaceAge predictions for non-cancerous clinical cohorts with predictions from the oncology datasets. We conducted all of the statistical analyses using Python (version 3.7) and R (version 3.6.3) with overall survival as a clinical endpoint. We used Kaplan–Meier curves to assess the model stratification power and the log-rank tests to test for differences between the stratified groups. Cox regression models were used to determine the impact of clinical covariates. We evaluated the model's explanatory power using the log-likelihood ratio test, concordance index, and area under the curve (AUC) of the receiver operating characteristic curve. In smaller datasets, we took different measures to prevent overfitting by selecting models that minimised the number of covariates and their SE while maximising concordance. Finally, to test real-world clinical use, we measured the effectiveness of FaceAge as a predictive metric against actual age by alternating substituting chronological age with FaceAge in the TEACHH model,[22] which is a clinically validated prognostic model that estimates life expectancy in patients with cancer undergoing palliative radiotherapy at the end of life. Additional details regarding the statistical and survival analyses are provided in the appendix (pp 19–20).

Testing clinical utility to improve physicians' end-of-life predictions

To compare FaceAge with human performance to predict the overall survival of patients with metastatic cancer, we performed a survey using 100 patients randomly sampled from the Harvard Palliative cohort. First, we assessed the performance of humans in estimating 6-month survival

from face photographs alone, without the benefit of additional clinical information, by asking ten medical and research staff members at Harvard-affiliated hospitals (five attending staff physicians who were all oncologists or palliative care physicians, three oncology residents, and two lay, non-clinical, researchers) to predict whether the patient would be alive after 6 months (an important endpoint to guide decision making at the end of life). We evaluated the performance of the ten survey-takers using the AUC and tested for statistically significant differences between the groups using a two-sided Wilcoxon signed-rank test. Furthermore, to evaluate the complementary value of FaceAge with other clinical data (ie, primary cancer diagnosis, age at treatment, performance status, location of metastases, number of emergency visits, number of hospital admissions, previous palliative chemotherapy courses, previous palliative radiotherapy courses, time to first metastasis, and time to oncology consult), we trained a FaceAge Risk Model, which combined the clinical factors with FaceAge to predict survival probability. In successive survey rounds, we asked the clinical survey-takers to predict 6-month survival based on the face photograph alone, the face photograph provided together with the patient clinical chart information, and then with the addition of the FaceAge Risk Model. Again, we evaluated the performance of the ten survey-takers using the AUC and tested for statistically significant differences between the groups using a two-sided Wilcoxon signed-rank test. More detailed information on the survey can be found in the appendix (pp 5, 29).

Genomic analysis

We evaluated the association of single-nucleotide polymorphisms (SNPs) with FaceAge or chronological age by running a gene-based analysis. Lymphocyte DNA from blood samples was collected, and whole-exome sequencing was conducted using the Illumina Infinium CoreExome Bead Chip (Illumina, San Diego, CA, USA). For quality control of the genotype data, we removed variants that violated Hardy–Weinberg equilibrium and had a missing rate of more than 5% and focused on variants with a minor allele frequency greater than 0·05. From the literature review,[23–32] we found the following known senescence genes: *TERT*, *ATM*, *CDKN1A*, *CDKN2B*, *TP53*, *IGFBP7*, and *MAPK10*. To identify a more inclusive network of senescence genes using a data-driven approach, we inputted the known senescence genes into GeneMania,[33] a platform that finds other genes related to those provided, to build a candidate network of senescence genes. GeneMania found a network of 27 genes (appendix p 11) and 22 of them had evaluable SNPs in our genotype dataset. We provide the rs IDs of these SNPs in the appendix (p 12). We ran the burden test using the GENESIS package[34] in R for the gene-based analysis. A gene-based analysis considers the aggregate effect of multiple variants in a test, and we used the burden test to perform this analysis.[35] The burden test collapses information of variants into a single genetic score,[36] and the association of this score with the outcome of interest, FaceAge or chronological age, is tested while adjusting for sex and population structure. To account for population structure, we applied a principal components analysis on the processed genotype data—ie, 145 283 variants that were autosomal, common (minor allele frequency >0·05), and linkage disequilibrium pruned (pairwise r^2 threshold of 0·5)—using PLINK 1.9, and included the first five principal components as covariates in the model. We provide the effect size with 95% CIs and score test p values for the genes.

Role of the funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Results

To assess the clinical relevance of FaceAge estimates in patients with a cancer diagnosis, we performed detailed experiments in three separate clinical cohorts from two institutions (figure 1). We first assessed the prognostic relevance of FaceAge predictions by their association with survival. The MAASTRO cohort included 4906 patients with a variety of non-metastatic cancer types and a wide range of prognoses (median age 66·0 years [range 22·0–94·0]; table 1). Kaplan–Meier survival analysis revealed good stratification of increasing mortality risk with increasing FaceAge risk groups (figure 2). All FaceAge risk groups showed significantly worse survival than the youngest-looking FaceAge risk group. This difference remained significant after adjusting for age, sex, and tumour group for the two oldest-looking risk groups (FaceAge >75 to 85 years and FaceAge >85 years) compared with the youngest-looking group (FaceAge ≤65 years). We observed similar results when assessing FaceAge as a continuous parameter, which showed significant prognostic performance (p=0·0013) after adjusting for age, sex, and tumour site in the whole cohort (figure 2; appendix p 14). Analysing specific cancer types, we found that FaceAge was significantly predictive in all cancer sites and remained significant after correcting for age and sex of patients with breast cancer, genitourinary cancer, and gastrointestinal cancer (figure 2; appendix p 14). In the MAASTRO cohort, adjusting for ethnicity was not possible, as this information was not collected in the electronic health record.

Next, we evaluated FaceAge in the Harvard Thoracic cohort, a site-specific dataset with thoracic malignancies (n=573; median age 69·0 years [range 33·3–93·2]; median actuarial overall survival 16·9 months; table 2), of which the majority were patients with non-small-cell lung cancer (n=450, 78·5%). Granular clinical data were available for these patients, allowing further investigation into the independent performance of FaceAge. Therefore, we investigated key clinical factors known to affect survival in lung cancer, including clinical stage, Eastern Cooperative Oncology Group (ECOG) performance status, smoking history, sex, histology, and treatment intent. In

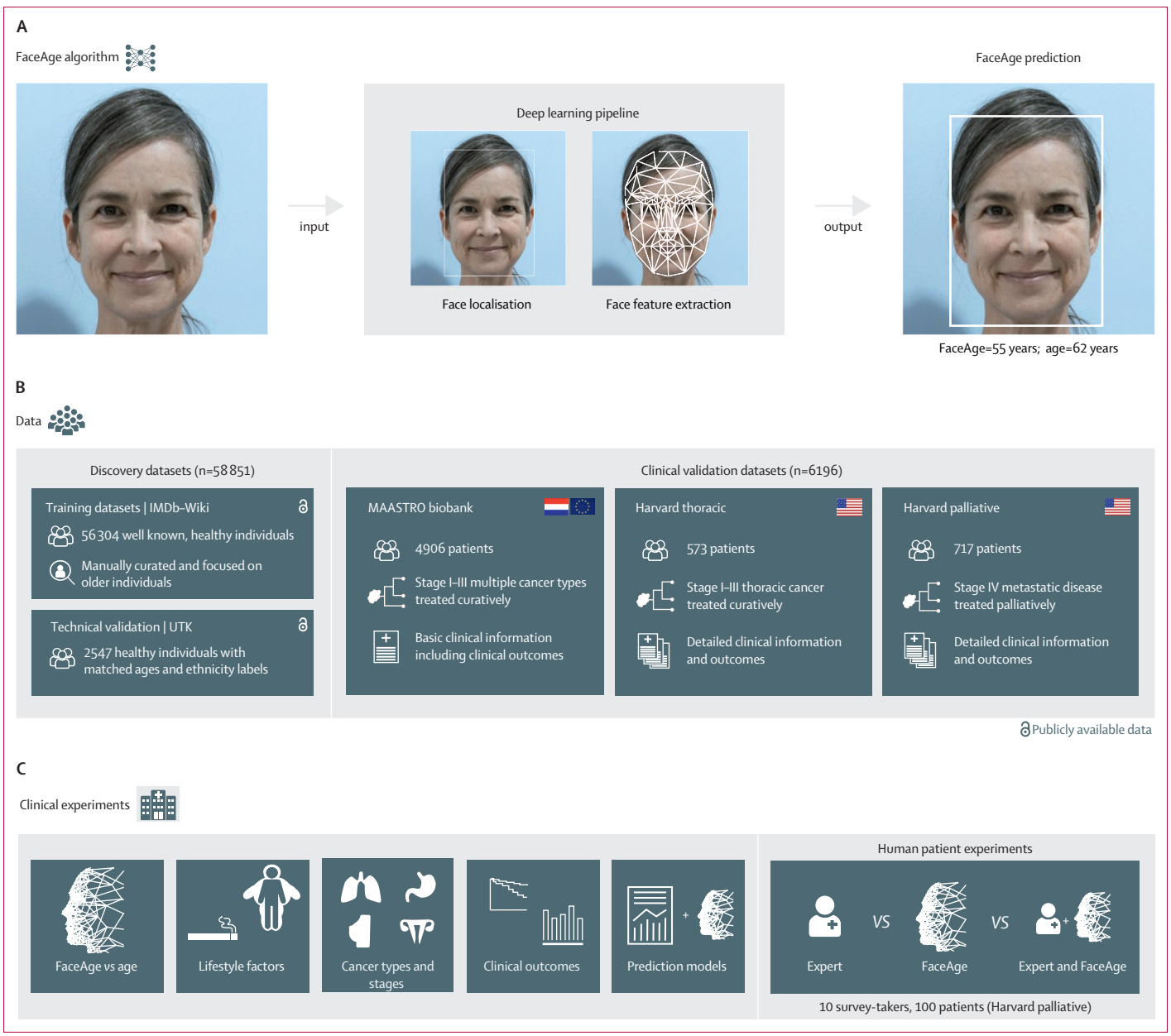

Figure 1: **Overview of the study method**
(A) The FaceAge algorithm uses a single photograph of the face as input. First, a convolutional neural network localises the face within the photograph. Next, a second convolutional neural network quantifies face features and uses these to predict the FaceAge of the person. (B) Overview of the datasets used in this study. The FaceAge algorithm was developed using a training dataset of 56 304 presumed healthy individuals (particularly politicians, actors, and other well known people). We assume that the people included in this cohort are of average health (ie, with a chronological age close to their biological age). This dataset was manually curated for individuals aged 60 years and older to enhance the dataset quality for the age range of the clinical oncology population. Three independent cohorts covering a large spectrum of patients with cancer were used to assess the clinical relevance of the algorithm. All patients had a face photograph acquired before radiation treatment as part of the routine clinical workflow. (C) Overview of the clinical experiments performed in this study to assess the clinical utility of FaceAge. Credits (face picture): RDNE Stock project from Pexels.

multivariable analysis, the association of FaceAge with overall survival was statistically significant after adjusting for these clinical factors (per decade hazard ratio [HR] 1·15 [95% CI 1·03–1·28]; p=0·011; appendix p 15). Comparing these results with those for age, we found that chronological age was not significantly associated with survival either on univariable analysis or after adjusting for multivariable clinical factors (1·08 [0·97–1·21]; p=0·16). Furthermore, we observed a significant increase in model explanatory power when adding FaceAge to the multivariate model (log-likelihood ratio test, χ^2 statistic, 1 degree of freedom 6·501; p=0·0108), whereas this was not observed when adding chronological age (1·965; p=0·16). These results show that FaceAge consistently improves

	Patients (N=4906)
Sex	
Male	2596 (52·9%)
Female	2310 (47·1%)
Age at treatment, years	66·0 (22·0–94·0)
Main cancer group	
Breast	1337 (27·3%)
Gastrointestinal	1003 (20·4%)
Genitourinary	843 (17·2%)
Lung	737 (15·0%)
Head and neck	456 (9·3%)
Other types of cancer	530 (10·8%)
Smoking history (patients for whom these data are available)	1302 (26·5%)
BMI (patients for whom these data are available)	1297 (26·4%)
ECOG performance status (patients for whom these data are available)	1170 (23·8%)

Data are n (%) or median (range). Ethnicity data were not collected for this dataset. ECOG=Eastern Cooperative Oncology Group.

Table 1: **Clinical characteristics of patients in the MAASTRO cohort**

prognostication whereas age does not, and that FaceAge contains prognostic information that is not captured by other investigated clinical parameters.

As a direct and relevant clinical application of FaceAge, we assessed the performance of FaceAge in patients at the end of life with metastatic cancer who received palliative treatment. In these patients, clinical prediction models can help to improve physicians' decision making as to whether or not to administer treatment, as well as the appropriate treatment intensity, both of which are largely a function of a physician's impression of overall prognosis, performance status, and frailty. We assessed the independent prognostic performance of FaceAge in the Harvard Palliative dataset (n=717; median age 65·2 years [range 19·6–97·6]; median actuarial overall survival 8·2 months; table 2). Covariates that are known to be related to survival in palliative patients with cancer,[22,37] such as performance status, number of hospital admissions and emergency room visits in the past 3 months, sites of metastatic disease, and the primary cancer type, were analysed. FaceAge was found to be significantly associated with or a significant predictor of survival in both univariable and multivariable analysis (univariable per decade HR 1·10 [95% CI 1·01–1·21]; p=0·035; multivariable per decade HR 1·12 [1·02–1·23]; p=0·021), whereas chronological age was not significant in either univariate or multivariate analysis (appendix pp 16–17). We also observed a significant increase in explanatory power by adding FaceAge to the multivariate model (log-likelihood ratio test, χ^2 statistic, 1 degree of freedom 5·439; p=0·020), which was not observed when adding chronological age (2·548; p=0·11).

Next, we evaluated the additive performance of FaceAge in the TEACHH model,[22] a clinically validated risk-scoring model for patients receiving palliative care. In the Harvard Palliative cohort, this model showed good performance in stratifying patients in different risk groups (appendix p 7). When substituting chronological age with FaceAge, we noticed a trend of increased log-likelihood ratio when comparing the three risk categories of the TEACHH model against the baseline hazard for FaceAge (log-likelihood ratio test, χ^2 statistic, 2 degrees of freedom 75·1; p<0·0001) compared with age (63·3; p<0·0001). This was also reflected in better separation of risk groups by survival, as substituting FaceAge lowered the median survival and increased the HR of the highest-risk group (FaceAge high-risk group median survival 0·21 years, HR 2·75; p<0·0001; chronological age high-risk group median survival 0·24 years, HR 2·43; p<0·0001), and raised median survival and decreased HR of the group at the lowest risk (2·2 years, HR 0·22; p<0·0001; 1·9 years, HR 0·28; p<0·0001).

Attending physicians performed the best overall, although there was a notable performance difference between individuals within each group (figure 3A). This was also shown by Kaplan–Meier analysis, in which the best-performing physician predicting 6-month survival was able to stratify patients into high-risk and low-risk groups that showed significant survival differences (median survival high-risk group 4·8 months *vs* low-risk group 13·2 months; p=0·0003), whereas the worst-performing physician did not (7·7 months *vs* 13·2 months; p=0·49; figure 3B).

Human performance significantly improved (p=0·0002) if we provided face photographs combined with clinical chart information (AUC 0·74 [95% CI 0·70–0·78]) compared with a face photograph only (0·61 [0·57–0·64]). However, human performance was improved even further (p<0·0001) when the FaceAge Risk Model was made available to clinicians in addition to chart information (AUC 0·80 [95% CI 0·76–0·83]), with the best performance of physicians not being statistically different (p=0·55) from the FaceAge risk model alone (0·81 [0·71–0·91]; figure 3C). Similar results were found for overall survival as quantified by the concordance index. We provide several case examples from the survey cohort to compare the 6-month survival predictions made by clinicians using different clinical aids with those of the TEACHH and FaceAge risk models (appendix p 10).

To evaluate whether FaceAge has the potential to be a biomarker for molecular ageing, we performed a gene-based analysis to measure its association with senescence genes in comparison with chronological age. The analysis was conducted on 146 individuals from the Harvard Thoracic Cohort who were diagnosed with non-small-cell lung cancer and profiled using whole-exome sequencing. We evaluated 22 genes known to be associated with senescence (appendix p 11), and we found that FaceAge was significantly associated with *CDK6* after adjusting for multiple comparisons (false discovery rate of 0·25; figure 3D). CDK6 has an important role in regulating the G1/S checkpoint of the cell cycle through phosphorylation and activation of the Rb (retinoblastoma) tumour suppressor protein by complexing with CDK4 and cyclin D. By contrast, no genes showed a significant association with chronological age after adjusting for multiple comparisons.

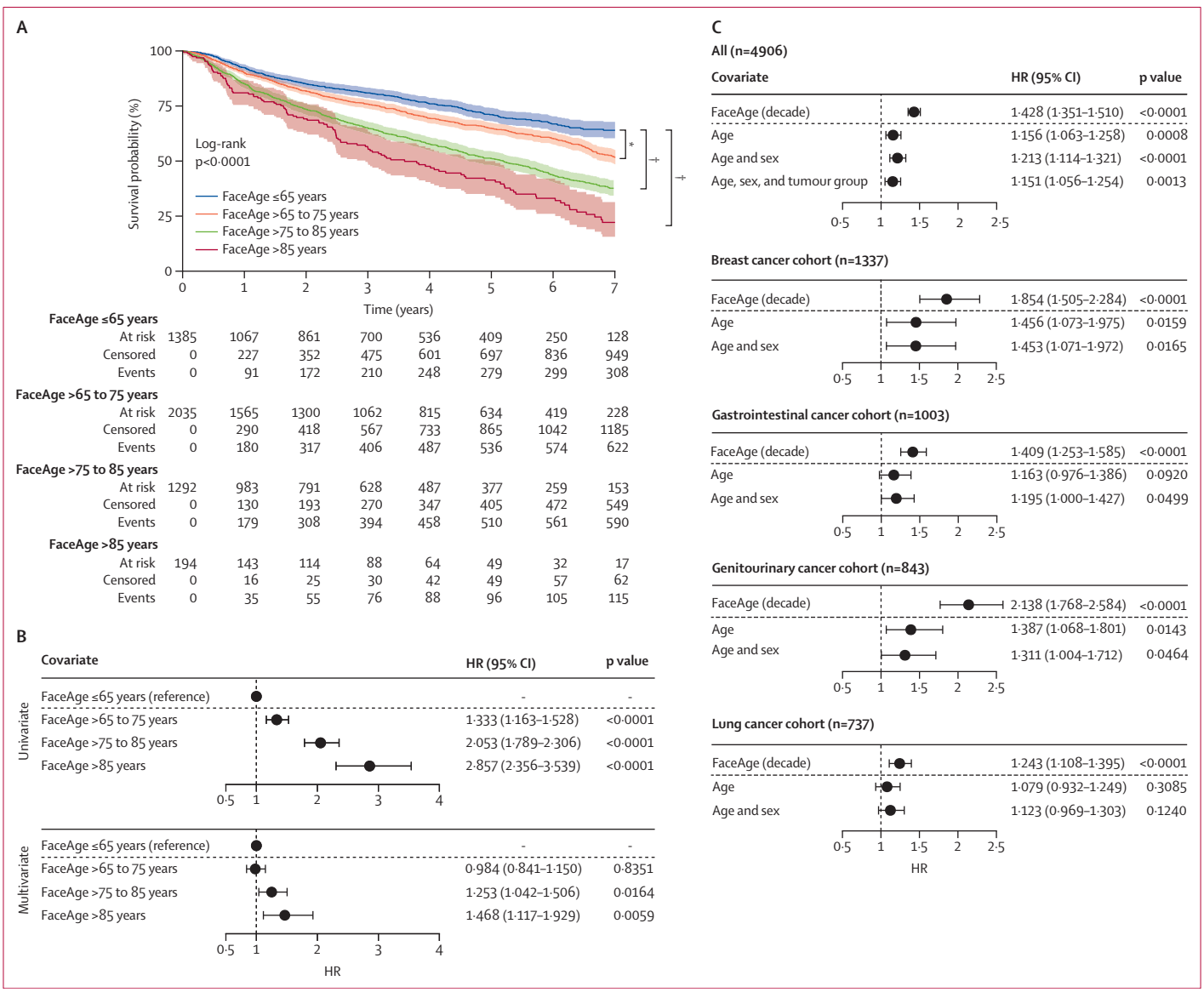

Figure 2: **Prognostic performance of FaceAge in several cancer cohorts**

(A) Kaplan–Meier survival analysis of FaceAge estimation with only a face photograph as input. Shaded areas are 95% CIs. (B) Forest plots of FaceAge risk groups. (C) Forest plots of FaceAge estimates as a continuous parameter for all patients as well as the four largest tumour sites. FaceAge is significant in all tumour sites in univariate analysis and remains significant in breast, gastrointestinal, and genitourinary cancer after correction for age and sex. All analyses are performed in the MAASTRO Cohort. HR=hazard ratio. *p<0·01. †p<0·001.

To evaluate the influence of cancer type and lifestyle factors on FaceAge predictions, we compared the difference between FaceAge and chronological age across cancer types, smoking history, BMI, and ECOG performance status (figure 4). We found that patients with cancer had a significantly higher FaceAge than chronological age (n=6367; mean increase of 4·79 years; paired two-sided *t* test p<0·001; figure 4A). This was consistent across cancer types and contrasted with the results in the presumed healthy populations. Firstly, in the UTK validation dataset with presumed healthy individuals, we found a significantly smaller difference between FaceAge and chronological age (mean increase of 0·35 years) compared with cancer cohorts (unpaired two-sided *t* test p<0·0001), indicating that individuals in the general population look more similar to their chronological age, as expected. Additionally, we analysed the faces of patients treated for benign conditions, as well as patients with ductal carcinoma in-situ. The non-cancerous cohorts had a smaller FaceAge-to-chronological age gap than the cohorts of patients with cancer (median difference 3·41 years *vs* 4·55 years for patients with cancer; p<0·0001), with the benign patients having a FaceAge closest to their chronological age (median difference 1·95 years compared with patients with cancer; p<0·0001), and patients with

	Patients
Harvard Thoracic cohort	(N=573)
Sex	
Male	270 (47·1%)
Female	303 (52·9%)
Age at treatment, years	69·0 (33·3–93·2)
Median overall survival, months	
Crude	14·4
Actuarial	16·9
Diagnosis	
Non-small-cell lung cancer	450 (78·5%)
Small-cell lung cancer	49 (8·6%)
Other types of cancer	74 (12·9%)
Treatment intent	
Curative non-stereotactic body radiation therapy	433 (75·6%)
Curative stereotactic body radiation therapy	106 (18·5%)
Palliative treatment	34 (5·9%)
Clinical stage	
I	145 (25·3%)
II	70 (12·2%)
III	279 (48·7%)
IV	70 (12·2%)
Not specified	9 (1·6%)
Tumour grade ≥2	322 (56·2%)
Ethnicity	
White	493 (86·0%)
African American	35 (6·1%)
Hispanic	16 (2·8%)
Asian	13 (2·3%)
Other	5 (0·9%)
Not available	11 (1·9%)
Smoking history (patients for whom these data are available)	509 (88·8%)
BMI (patients for whom these data are available)	106 (18·5%)
ECOG performance status ≥2	109 (19·0%)
Harvard Palliative cohort	(N=717)
Sex	
Male	333 (46·4%)
Female	384 (53·6%)
Ethnicity	
White	605 (84·4%)
Black	42 (5·9%)
Asian	28 (3·9%)
Latino	5 (0·7%)
Other	13 (1·8%)
Not available	24 (3·3%)
Age at treatment, years	65·2 (19·6–97·6)
Median overall survival, months	
Crude	4·5
Actuarial	8·2
Time from diagnosis to metastasis, months	1·0 (0–26·8)
Time from metastasis to radiotherapy consultation, months	4·5 (0–33·0)
Primary cancer diagnosis	
Lung	201 (28·0%)
Breast	118 (16·5%)
Prostate	58 (8·1%)
Colorectal	44 (6·1%)
Gynaecological	43 (6·0%)
Melanoma	38 (5·3%)
Oesophagus	24 (3·3%)
Renal	24 (3·3%)
Sarcoma	24 (3·3%)
Unknown primary	20 (2·8%)
Bladder	19 (2·6%)
Head and neck	10 (1·4%)
Pancreas	18 (2·5%)
Neuroendocrine	12 (1·7%)
Cholangiocarcinoma	9 (1·3%)
Hepatocellular	8 (1·1%)
Non-melanoma skin	8 (1·1%)
Stomach	8 (1·1%)
Genitourinary (non-bladder or testicular)	6 (0·8%)
Mesothelioma	5 (0·7%)
Testicular	4 (0·6%)
Small bowel	3 (0·4%)
Other	6 (0·8%)
Metastases	
Bone	344 (48·0%)
Brain	265 (37·0%)
Spine	256 (35·7%)
Lung	223 (31·1%)
Liver	222 (31·0%)
Lymph	219 (30·5%)
Adrenal	65 (9·1%)
Other	153 (21·3%)
ECOG performance status	
0–1	309 (43·1%)
2	135 (18·8%)
3	110 (15·3%)
4	11 (1·5%)
Previous palliative radiotherapy, courses	
0	627 (87·4%)
1	72 (10·0%)
≥2	14 (2·0%)
Previous palliative chemotherapy, courses	
0	280 (39·1%)
1	197 (27·5%)
≥2	237 (33·1%)
Hospital admissions	
0	370 (51·6%)
1	291 (40·6%)
≥2	55 (7·7%)
Emergency room visits	
0	421 (58·7%)
1	233 (32·5%)
≥2	62 (8·6%)

(Table 2 continues in next column)

(Continued from previous page)

Data are n (%) or median (range). ECOG=Eastern Cooperative Oncology Group.

Table 2: **Clinical characteristics of patients in the Harvard cohorts**

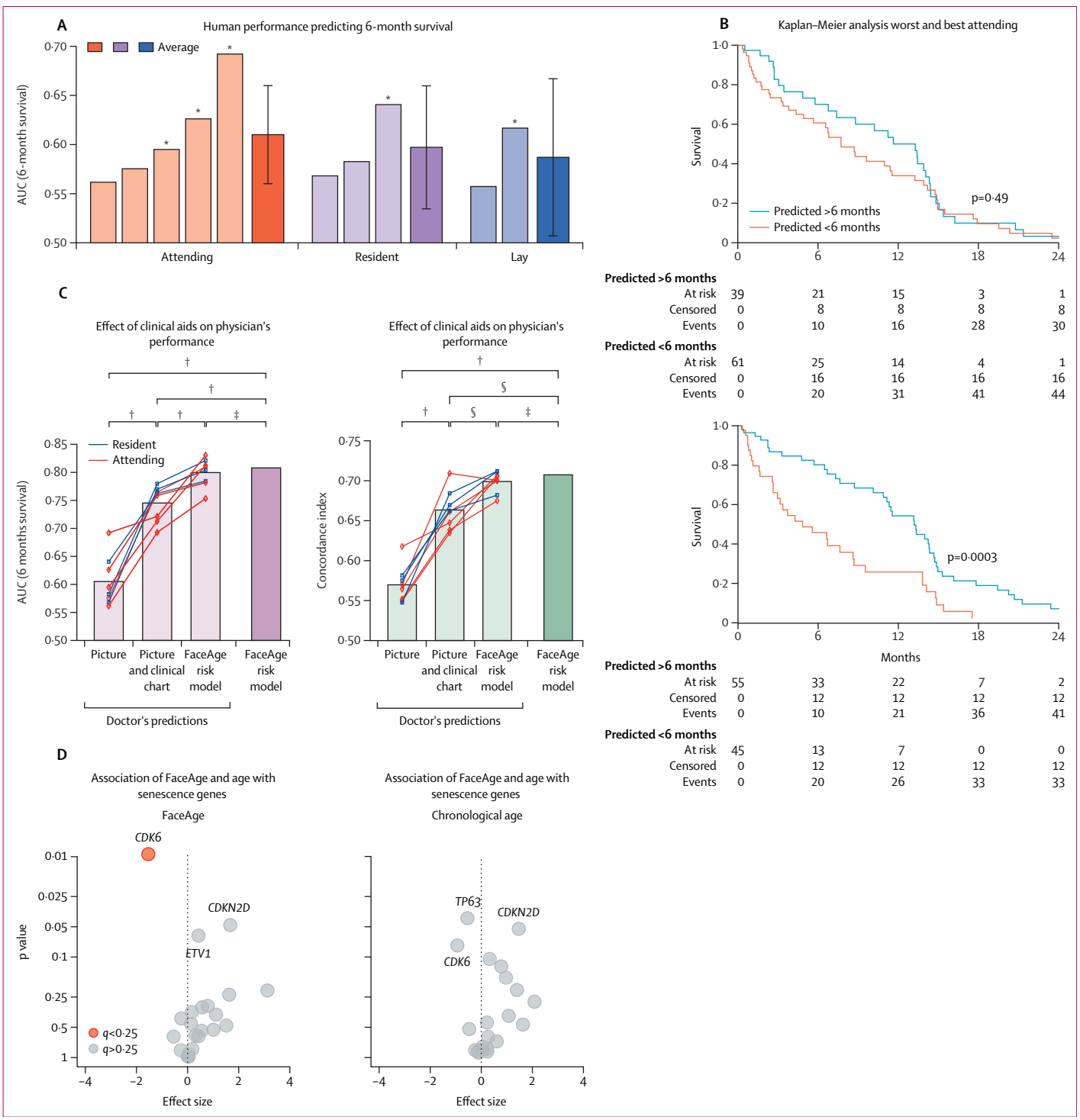

Figure 3: **Comparison of human and FaceAge performance predicting survival**

(A) AUC of the receiver operating characteristic for 6-month survival predicted by ten survey-takers, grouped by experience level: attending physicians (oncologists or palliative care physicians), oncology residents, and lay researchers (non-clinical). CIs are shown for average AUCs. (B) Kaplan–Meier analysis of overall survival of patients predicted to be either alive or not at 6 months, comparing the lowest (top graph) and highest (bottom graph) performers of the attending physicians. (C) 6-month survival prediction (left graph) and overall survival time (right graph) for physicians (both attending and residents) aided with only a picture, a picture and clinical chart information, and a risk model including clinical data and FaceAge. The survey included 100 patients receiving palliative care who were randomly selected from the Harvard Palliative cohort. (D) Results of the burden test for the association of the senescence genes with FaceAge or chronological age. After adjusting for multiple comparisons using a false discovery rate (*q*) of 0·25, only *CDK6* was statistically significant for FaceAge, while none of the other genes were statistically significant for either FaceAge or chronological age. AUC=area under the curve. *AUC significantly different from random. †p<0·01. ‡Not significant. §p<0·05.

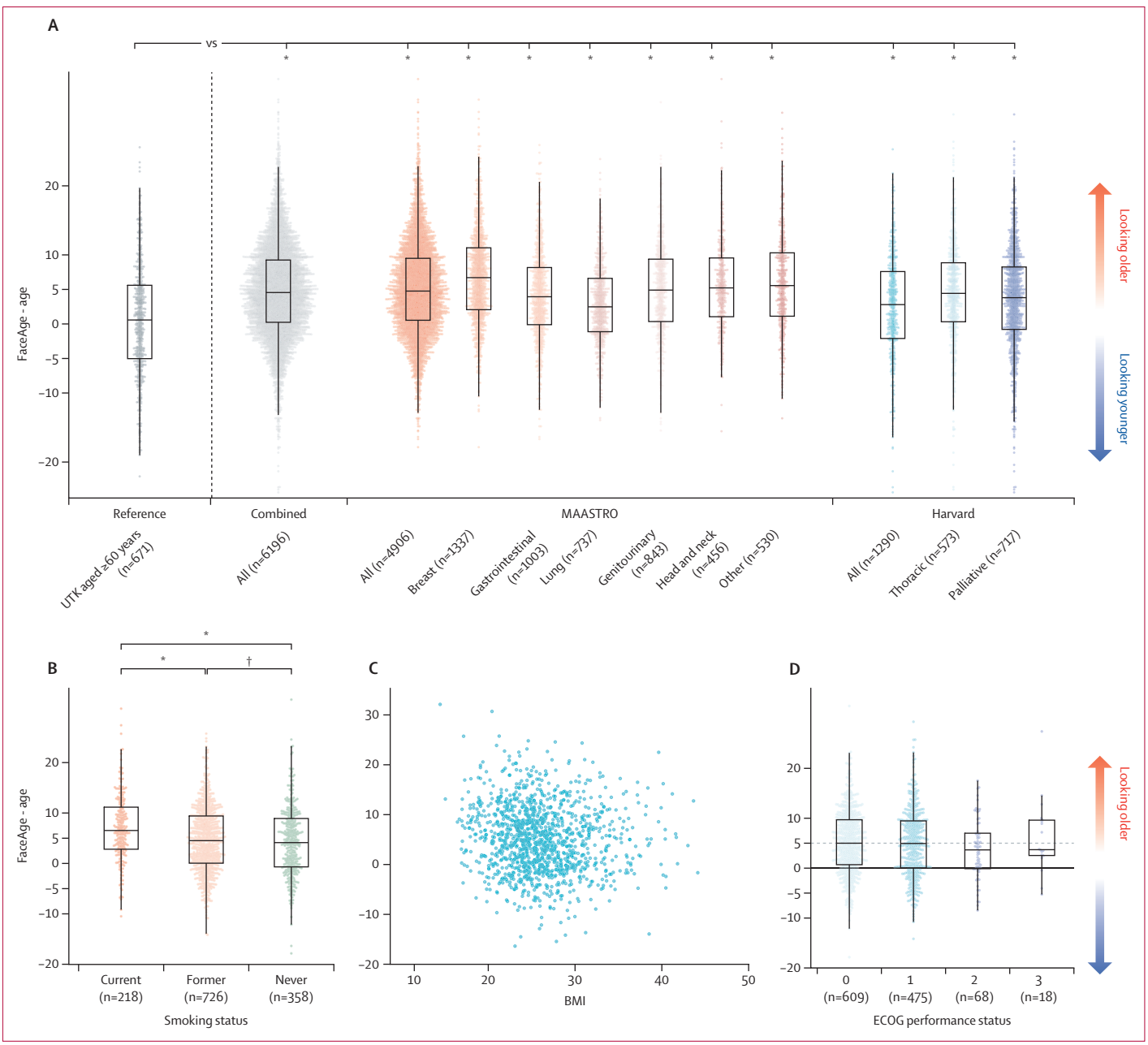

Figure 4: **Application of FaceAge in patients with cancer**

(A) Difference between FaceAge and age across cancer types and datasets to investigate whether individuals look older or younger than their chronological age. (B) Difference between FaceAge and age for current, former, and never smokers included in the MAASTRO cohort. (C) Scatterplot assessing the association of FaceAge with BMI in the MAASTRO cohort. (D) Association of FaceAge with Eastern Cooperative Oncology Group performance status was quantified for a subset of patients in the MAASTRO cohort. In the boxplots, the box shows the IQR, with the line at the centre of the box indicating the mean; the top (bottom) whisker extends from the box to the largest (smallest) value within 1·5 IQR. An unpaired, two-sided *t* test was used in all cases. ECOG=Eastern Cooperative Oncology Group. *p<0·001. †Not significant.

ductal carcinoma in situ having intermediate FaceAge values (median difference 3·86 years compared with patients with cancer; p=0·019; appendix p 8).

To assess the effect of lifestyle factors, we compared the difference between FaceAge and chronological age in current, former, and never smokers in the MAASTRO cohort. We found that current smokers look significantly older (mean increase 33·24 months; unpaired two-sided *t* test *t*=4·78 [95% CI 1·63–3·91]; p<0·001) than former and never smokers (figure 4B), which was consistent across cancer types (appendix p 9). In an assessment of the effect of BMI on the difference between FaceAge and chronological age (figure 4C; appendix p 9), we observed a statistically significant association (n=1295; *r* –0·0999; p<0·0001), but the

effect size was minimal, indicating a weak relationship between FaceAge and BMI. Since the ECOG performance status is used for clinical stratification, we compared the association of ECOG groups with the difference between FaceAge and chronological age (figure 4D; appendix p 9). In both the MAASTRO and Harvard cohorts, we found no statistically significant differences (unpaired two-sided *t*-test p>0·092) between the groups, indicating that FaceAge quantifies biological information that differs from the performance status of a patient.

Discussion

In this study, we showed that facial features captured on easily obtainable face photographs contain prognostic information related to the apparent age of a person. Our method, which relies on deep learning to automatically extract these facial features, was used to develop a new biomarker, FaceAge, that was able to stratify a broad spectrum of patients with cancer according to survival risk. FaceAge was found to be more predictive than actual (chronological) age in independent, heterogeneous datasets and to improve on the standard of subjective visual assessment by clinicians. We focused our work on patients treated with radiotherapy since they are closely followed regarding their survival outcomes, and their disease processes and treatments can substantially affect their biological ageing. Furthermore, unlike other patients, patients with cancer routinely have their face photographs taken as part of the treatment registration process.

We found that, on average, patients with cancer look approximately 5 years older than their chronological age and have a statistically higher FaceAge compared with clinical cohorts of patients without cancer who are treated for conditions that are benign or precancerous. FaceAge outperformed age in univariate and multivariate analyses across several cancer sites and clinical subgroups, even after adjusting for known clinical risk factors, and provided additional explanatory power. Notably, FaceAge performed well in patients treated for curative intent, with life expectancies of several years, and in patients at the end of life with an expected survival of weeks to months. To test how clinicians might use FaceAge, we also showed that FaceAge significantly improved the performance of a validated clinical risk-scoring model[27] for estimating survival in patients at the end of life who received palliative radiation treatment, a patient population for which improvements in treatment decision making using such models are critical. We showed that survival prediction performance of clinicians improved when FaceAge risk model predictions were made available, especially among physicians with lower baseline performance. Lastly, we provided evidence from SNP gene analysis that FaceAge is correlated with molecular processes of cell-cycle regulation and cellular senescence, supporting the hypothesis that FaceAge is a biomarker that relates to biological ageing, consistent with its interpretation as a modifier of survival time in a diseased population. Multiple publications have discussed evidence of the *CDK6* gene delaying senescence,[38,39] which agrees with the inverse association that we observed between *CDK6* and FaceAge. However, since our genomic analysis was on 146 patients with non-small-cell lung cancer from the Harvard Thoracic cohort, our results could also be attributed to the involvement of the *CDK6* gene in cancer. For this study, we did not have access to a healthy cohort to quantify such a relation. In future investigations, having a healthy control group could help to confirm that the correlation we observed is not entirely related to cancer. Nevertheless, although limited in scope to only a small set of preselected genes to conserve statistical power, our analysis shows the potential of using FaceAge to discover associations with genes related to biological ageing, which are different from and might not be detected by chronological age.

Our pipeline was entirely trained on large, publicly available, non-clinical databases. Although the standard approach in deep learning is to train the models on datasets very similar to the test population, the performance of the FaceAge model explicitly relies on a presumed difference between healthy populations and populations with health conditions, with the hypothesis that the predicted age differential reflects a component unrelated to model error but is instead attributable to the intrinsic difference between age and biological age. The training data we used might have limitations and possible intrinsic biases. As the online image databases used for training do not include associated health information, we made the implicit assumption that the patients in the training data were of average health for their age (ie, they have a biological age similar to their chronological age), although this assumption is clearly not true in all cases. Moreover, the images contain a substantial proportion of well known individuals, such as actors and politicians, which might introduce a systematic biological age selection bias as such individuals might have a different biological age compared with an age-matched cohort of non-famous peers, due to different lifestyle and socioeconomic factors. For example, actors and other individuals in the public eye might have cosmetic or other facial alterations that could affect biological age estimations from such photographs, in addition to more frequent digital image touch-ups. Although we do not have statistics on how many of the photographs included cosmetic alterations of the physical or digital variety, this is unlikely to represent a substantial proportion of the whole, as the IMDb–Wiki database contains many photos of people other than actors, including writers, philanthropists, educators, scientists, and people from all domains of society. Moreover, the large size of the heterogeneous dataset tends to average out potential biasing factors. Notably, in some of the smaller cancer cohorts, after adjusting for various clinical covariates, we noticed a weakening of the association of FaceAge with survival. We hypothesise that the reduced sample sizes and the residual age imbalance between the training data and the clinical test datasets could have contributed to such a loss of significance. Further validation in larger cohorts is warranted to verify our results.

The use of facial photographs in our analyses presents multiple ethical considerations. Outside of health care, we acknowledge the potential for misuses of a model trained on such data. Examples include the incorporation by health, disability, and life insurance payors of estimated survival metrics from face images to determine the insurability of prospective policy holders, or the promotion by technology or media companies of health or lifestyle products with targeted advertising based on client biological age estimation. Strong regulatory oversight would be a first measure towards mitigating this problem. Another important ethical concern is racial or ethnic bias, which has been problematic for automated face recognition software, especially in legal and law enforcement applications. In our work, we studied the potential for racial bias in several ways: first, we quantified the model age predictions across different ethnic groups drawn from the UTK validation dataset, finding that the model is affected, although not substantially, by the patients' ethnicity (appendix p 13). The UTK is one of the most ethnically diverse age-labelled face image databases available publicly and, therefore, appropriate for assessing model performance in this regard, with non-White individuals comprising approximately 55% of the database. Second, ethnicity was treated as a covariate that we adjusted for in the multivariable analysis of the Harvard clinical datasets, which revealed again that the FaceAge measure was minimally affected by ethnicity. This could not be done in the MAASTRO cohort, since ethnicity data are not routinely collected at this centre. Our model is configured for the task of age estimation, which, in our opinion, has less embedded societal bias than the task of face recognition. However, further assessments of bias in performance across different populations will be essential, as differential treatment decisions as a result of its predictions could amplify existing health disparities—and more research is needed on this topic. In particular, the acquisition of new, more diverse, and representative datasets (in both clinical and non-clinical settings) will be a key step to better address bias during training or to study it in more detail during validation. Institutional and governmental oversight of how such models are regulated and deployed, with careful prescription of their intended use and educational support for clinical end-users (including appropriate use case and model failure modes), will be crucial to ensure that patients can benefit from their incorporation into clinical care while minimising the risk of abuse, unintended or otherwise. Before clinical implementation, further work is needed to address these technical and ethical concerns, including: optimisation and standardisation of training datasets to account for potential technical, health-related, and racial biases; validation in clinical datasets representative of the target population for a given clinical use case; and additional correlation with features or molecular markers of biological ageing.

In conclusion, our results suggest that a deep learning model can enhance survival prediction in patients with cancer through the analysis of face photographs and estimation of patients' biological age from their facial features. We have shown that deep learning-based FaceAge estimates might be prognostic in a wide range of cancer types and clinical settings, and that integration with existing clinical chart information and clinical risk-scoring models might improve clinicians' prediction performance. However, further research and development must be carried out before this technology can be effectively deployed in a real-world clinical setting.

For the **IMDb–Wiki dataset** see https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki

For the **UTK-face dataset** see https://susanqq.github.io/UTKFace

Contributors

The study was conceptualised by HJWLA, DB, RHM, and OZ. The model design and implementation were carried out by DB and OZ. The data curation, analysis, and interpretation were done by HJWLA, DSB, DB, AD, RHM, DDR, and OZ. The statistical analyses were done by HJWLA, DB, and OZ. The study was supervised by HJWLA and RHM. The data were directly accessed and verified by HJWLA, DB, DSB, RHM, and OZ. The initial draft of the manuscript was written by HJWLA, DB, RHM, and OZ. All authors subsequently read and revised the initial draft and read and approved the final version. All authors were frequently updated about results and their interpretations were used to improve the study. All authors had access to all the data in the study and have final responsibility for the decision to submit for publication.

Declaration of interests

Mass General Brigham submitted a patent partly based on the results presented in this study to the US Patent and Trade Office. DSB reports honoraria from the American Society of Clinical Oncology and Harvard PGME. AD reports grants or contracts from the European Commission, Janssen Cilag, IQVIA, ZonMw, The Dutch Research Council, Health Holland, The Dutch National Growth Fund, Mirada Medical, and Varian and Siemens Healthineers; payments or honoraria from Janssen Cilag, Medtronic, and Roche; financial support from IQVIA; serving in a leadership role with the Hanarth Fund, the MD Anderson Cancer Center (Houston, TX, USA), the Peter Munk Cardiac Centre (Toronto, ON, Canada), the Novo Nordisk Foundation (Copenhagen, Denmark), and the Luxemburg National Research Fund (Esch-sur-Alzette, Luxembourg); and stock options in Medical Data Works. DDR reports grants or contracts from AstraZeneca, Bristol Myers Squibb, BeiGene, Philips, Olink, and Eli Lilly. BH-K reports receiving grant or contracts from the Botha-Chan Low Grade Glioma Consortium. CS reports support from the Francis Crick Institute and the Royal Society; grants or contracts from AstraZeneca, Boehringer Ingelheim, Bristol Myers Squibb, Invitae (formerly Archer Dx), Ono Pharmaceuticals, Pfizer, and Roche-Ventana; consulting fees from Bicycle Therapeutics, Genentech, Medicxi, Metabomed, and Novartis; being a member of GRAIL, Relay Therapeutics, the China Innovation Centre of Roche, SAGA Diagnostics, and the Sarah Cannon Research Institute; honoraria from Amgen, AstraZeneca, Bristol Myers Squibb, Illumina, GSK, MSD, Roche-Ventana, and Pfizer; patents planned, issued, or pending, including PCT/GB2017/053289, PCT/EP2016/059401, PCT/EP2016/071471, PCT/GB2018/052004, PCT/GB2020/050221, PCT/GB2018/051912, PCT/US2017/28013, and PCT/GB2018/051892; leadership or fiduciary roles with Cancer Research UK and the American Association for Cancer Research; stock or stock options in Apogen Biotech, Epic Biosciences, GRAIL, and Achilles Therapeutics; and other financial or non-financial interests with AstraZeneca and GRAIL Bio UK. OZ received support from Queen's University (Kingston, ON, Canada). HJWLA reports consulting fees or stock from Onc AI, Love Health, AstraZeneca, Health-AI, Ambient, and Sphera, outside the submitted work. All other authors declare no competing interests.

Data sharing

The following datasets are publicly available and can be downloaded directly from the respective websites: IMDb–Wiki dataset, a large dataset of face photographs of individuals included in IMDb and Wikipedia that was used to develop the FaceAge model, and the UTKFace dataset, an independent dataset that was used for the technical validation of our model. No clinical

datasets can be shared owing to institutional research ethics board protocols and privacy concerns regarding face photographs of patients. The output data, including artificial intelligence-predicted ages in the UTK dataset, are provided online. The code is provided at the Github FaceAge repository.

Acknowledgments

We acknowledge financial support from the National Institutes of Health (U24CA194354, U01CA190234, U01CA209414, R35CA22052, and U54CA27451 to HJWLA; K08DE030216-01 to BHK), and the EU European Research Council (866504 to HJWLA).

References

1 Bell CG, Lowe R, Adams PD, et al. DNA methylation aging clocks: challenges and recommendations. *Genome Biol* 2019; **20:** 249.
2 Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. *Aging Cell* 2015; **14:** 924–32.
3 Xie W, Baylin SB, Easwaran H. DNA methylation in senescence, aging and cancer. *Oncoscience* 2019; **6:** 291–93.
4 Córdova-Oriz I, Chico-Sordo L, Varela E. Telomeres, aging, and reproduction. *Curr Opin Obstet Gynecol* 2022; **34:** 151–58.
5 Blackburn EH, Greider CW, Szostak JW. Telomeres and telomerase: the path from maize, *Tetrahymena*, and yeast to human cancer and aging. *Nat Med* 2006; **12:** 1133–38.
6 Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB. Telomere end-replication problem and cell aging. *J Mol Biol* 1992; **225:** 951–60.
7 Weindruch R, Kayo T, Lee CK, Prolla TA. Gene expression profiling of aging using DNA microarrays. *Mech Ageing Dev* 2002; **123:** 177–93.
8 McCarroll SA, Murphy CT, Zou S, et al. Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat Genet* 2004; **36:** 197–204.
9 Karasik D, Hannan MT, Cupples LA, Felson DT, Kiel DP. Genetic contribution to biological aging: the Framingham Study. *J Gerontol A Biol Sci Med Sci* 2004; **59:** 218–26.
10 Jain R, Duval S, Adabag S. How accurate is the eyeball test? A comparison of physician's subjective assessment versus statistical methods in estimating mortality risk after cardiac surgery. *Circ Cardiovasc Qual Outcomes* 2014; **7:** 151–56.
11 Glare P, Virik K, Jones M, et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ* 2003; **327:** 195–98.
12 Cheon S, Agarwal A, Popovic M, et al. The accuracy of clinicians' predictions of survival in advanced cancer: a review. *Ann Palliat Med* 2016; **5:** 22–29.
13 White N, Reid F, Harris A, Harries P, Stone P. A systematic review of predictions of survival in palliative care: how accurate are clinicians and who are the experts? *PLoS One* 2016; **11:** e0161407.
14 Bobrov E, Georgievskaya A, Kiselev K, et al. PhotoAgeClock: deep learning algorithms for development of non-invasive visual biomarkers of aging. *Aging (Albany NY)* 2018; **10:** 3249–59.
15 Clapes A, Anbarjafari G, Bilici O, Temirova D, Avots E, Escalera S. From apparent to real age: gender, age, ethnic, makeup, and expression bias analysis in real age estimation. 2018. https://ieeexplore.ieee.org/document/8575487 (accessed Nov 20, 2023).
16 Punyani P, Gupta R, Kumar A. Neural networks for facial age estimation: a survey on recent advances. *Artif Intell Rev* 2020; **53:** 3299–347.
17 Rothe R, Timofte R, Van Gool L. Deep expectation of real and apparent age from a single image without facial landmarks. *Int J Comput Vis* 2018; **126:** 144–57.
18 Xia X, Chen X, Wu G, et al. Three-dimensional facial-image analysis to predict heterogeneity of the human ageing rate and the impact of lifestyle. *Nat Metab* 2020; **2:** 946–57.
19 Zhang Z, Song Y, Qi H. Age progression/regression by conditional adversarial autoencoder. 2017. http://openaccess.thecvf.com/content_cvpr_2017/html/Zhang_Age_ProgressionRegression_by_CVPR_2017_paper.html (accessed Nov 20, 2023).
20 Zhang K, Zhang Z, Li Z, Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process Lett* 2016; **23:** 1499–503.
21 Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, Inception-ResNet and the impact of residual connections on learning. *AAAI* 2017; 4278–84.
22 Krishnan MS, Epstein-Peterson Z, Chen Y-H, et al. Predicting life expectancy in patients with metastatic cancer receiving palliative radiotherapy: the TEACHH model. *Cancer* 2014; **120:** 134–41.
23 van Moorsel CHM. Trade-offs in aging lung diseases: a review on shared but opposite genetic risk variants in idiopathic pulmonary fibrosis, lung cancer and chronic obstructive pulmonary disease. *Curr Opin Pulm Med* 2018; **24:** 309–17.
24 Dressen A, Abbas AR, Cabanski C, et al. Analysis of protein-altering variants in telomerase genes and their association with *MUC5B* common variant status in patients with idiopathic pulmonary fibrosis: a candidate gene sequencing study. *Lancet Respir Med* 2018; **6:** 603–14.
25 Sclafani A, Worsham CM, McNeill J, Anandaiah A. Advances in interstitial lung disease genetics. *Am J Respir Crit Care Med* 2019; **200:** 247–49.
26 Newton CA, Oldham JM, Ley B, et al. Telomere length and genetic variant associations with interstitial lung disease progression and survival. *Eur Respir J* 2019; **53:** 1801641.
27 Arimura-Omori M, Kiyohara C, Yanagihara T, et al. Association between telomere-related polymorphisms and the risk of IPF and COPD as a precursor lesion of lung cancer: findings from the Fukuoka Tobacco-Related Lung Disease (FOLD) Registry. *Asian Pac J Cancer Prev* 2020; **21:** 667–73.
28 Bouten RM, Dalgard CL, Soltis AR, Slaven JE, Day RM. Transcriptomic profiling and pathway analysis of cultured human lung microvascular endothelial cells following ionizing radiation exposure. *Sci Rep* 2021; **11:** 24214.
29 Liu X, Shao C, Fu J. Promising biomarkers of radiation-induced lung injury: a review. *Biomedicines* 2021; **9:** 1181.
30 Chen H, Chen H, Liang J, et al. TGF-β1/IL-11/MEK/ERK signaling mediates senescence-associated pulmonary fibrosis in a stress-induced premature senescence model of Bmi-1 deficiency. *Exp Mol Med* 2020; **52:** 130–51.
31 Picard D, Echeverria PC. HSP90 chaperone cycle for steroid hormone receptors (SHR). 2017. https://reactome.org/content/detail/R-HSA-3371497.1 (accessed Nov 20, 2023).
32 Gene Ontology Consortium. AmiGO 2. Mitogen-activated protein kinase 10. https://amigo.geneontology.org/amigo/gene_product/UniProtKB:P53779 (accessed Nov 4, 2023).
33 Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 2010; **38** (suppl 2): W214–20.
34 Gogarten SM, Sofer T, Chen H, et al. Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics* 2019; **35:** 5346–48.
35 Chung J, Jun GR, Dupuis J, Farrer LA. Comparison of methods for multivariate gene-based association tests for complex diseases using common variants. *Eur J Hum Genet* 2019; **27:** 811–23.
36 Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 2014; **95:** 5–23.
37 Krishnan M, Temel JS, Wright AA, Bernacki R, Selvaggi K, Balboni T. Predicting life expectancy in patients with advanced incurable cancer: a review. *J Support Oncol* 2013; **11:** 68–74.
38 Wagner V, Gil J. Senescence as a therapeutically relevant response to CDK4/6 inhibitors. *Oncogene* 2020; **39:** 5165–76.
39 Ruas M, Gregory F, Jones R, et al. CDK4 and CDK6 delay senescence by kinase-dependent and $p16^{INK4a}$-independent mechanisms. *Mol Cell Biol* 2007; **27:** 4273–82.

For the **output data** see https://aim.hms.harvard.edu/faceage

For the **Github FaceAge repository** see https://github.com/AIM-Harvard/FaceAge