FACEAGE IN PATIENTS WITH OLIGOMETASTATIC CANCER: EXTERNAL VALIDATION AND RACIAL DIFFERENCES IN FACIAL AGE ESTIMATES

Anonymous authors
Paper under double-blind review

ABSTRACT

Purpose: FaceAge is a deep learning biomarker that estimates biological age from facial photographs and has been associated with survival in patients with metastatic cancer. We hypothesized that FaceAge would predict survival in patients with oligometastases and that the newer foundation model FaceAge 2.0 would have smaller racial differences compared to FaceAge 1.0. Methods & Materials: Between 2014 and 2023, 177 consecutive patients with 1 to 5 distant metastases were treated with comprehensive involved-site radiation at a community hospital. Frontal facial photographs obtained during radiation simulation were processed with FaceAge 1.0 and FaceAge 2.0/FAHR-Face. Photographs from 167 patients met quality criteria. Associations with overall survival (OS) were evaluated using Kaplan–Meier estimates, Cox regression and the concordance index. We compared FaceAge estimates across racial groups after adjusting for chronological age. Results: At a median follow-up of 44.0 months, the median OS was 42.8 months. After adjusting for chronological age, FaceAge 1.0 scores were underestimated in Black patients, with a mean difference of 10.6 years ($p < 0.001$) compared to White patients. When using FaceAge 2.0, the difference decreased to 3.0 years ($p = 0.16$). On univariable analysis, FaceAge 1.0 (HR 1.03 per year, $p = 0.004$) and FaceAge 2.0 (HR 1.03 per year, $p = 0.002$) both predicted OS, consistent with previous FaceAge studies. Chronological age was also strongly prognostic (HR 1.04 per year, $p < 0.001$) and remained significant on multivariable analysis. Adding FaceAge to chronological age and established prognostic factors only modestly improved discrimination (C-index 0.74 without FaceAge vs. 0.75 with FaceAge 1.0 or 2.0). Conclusions: FaceAge was externally validated as a univariable prognostic biomarker in oligometastatic cancer. FaceAge 2.0 substantially reduced racial disparities in facial age estimates seen with FaceAge 1.0. Further studies are warranted to determine whether facial aging models perform consistently across racial groups before routine clinical use.

Keywords: algorithmic fairness, facial age estimation, computer vision, foundation models, medical artificial intelligence, oncology

1 INTRODUCTION

Chronological age is a key prognostic factor in clinical oncology that is routinely used to guide treatment decisions (Magnuson et al., 2024; Dale et al., 2023). Practicing oncologists recognize that chronological age alone incompletely captures frailty, physiologic reserve, functional status, comorbidity, or treatment tolerance (Dale et al., 2023). In a prospective trial, survival predictions based on an experienced nurse's subjective gestalt assessment performed comparably with physician predictions based on more granular clinical and genomic variables (Kao et al., 2022). Accurately capturing this previously unquantified data is particularly relevant in oligometastatic cancer because a significant minority of treated patients can achieve durable disease control and survive long enough for competing mortality and aging biology to impact overall survival (Kao et al., 2023).

Advances in deep learning and facial recognition technology have enabled the development of algorithms that objectively quantify visual indicators of biological aging and overall health (Bobrov

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

et al., 2018). FaceAge is an artificial intelligence (AI) algorithm used to estimate biological age based on a single digital face photograph. In prior cancer cohorts, FaceAge was strongly correlated with overall survival, indicating that facial appearance captures clinically meaningful aging information not fully reflected by chronological age alone (Bontempi et al., 2025; Lee et al., 2026). Since radiation oncology clinics routinely obtain simulation photographs for patient identification, FaceAge is a novel and promising biomarker that does not add additional cost or patient burden.

Accurately predicting outcome for patients with oligometastases remains an unmet clinical need (Palma et al., 2015; Kao et al., 2024). AI-derived biomarkers require external validation across clinical settings and demographic groups before clinical implementation. The original FaceAge 1.0 model was developed using more limited training data, while FaceAge 2.0/FAHR-Face uses a large foundation-model approach that may improve generalizability and fairness across demographic groups (Haugg et al., 2025). We externally validated the prognostic performance of FaceAge 1.0 and FaceAge 2.0 in oligometastatic cancer and evaluated whether FaceAge 2.0 reduced racial bias compared with FaceAge 1.0.

2 Materials and Methods

2.1 Patients and Curated Data

This retrospective study was approved by the institutional review board at the participating institution. The dataset consists of 177 consecutive patients with oligometastases, defined as 1 to 5 distant metastases, treated with comprehensive involved-site radiation from 2014 to 2023 by a single radiation oncologist.

Clinical variables recorded included chronological age, race, sex, ECOG performance status, primary tumor site, the number of distant metastatic lesions, EORTC metastasis classification (synchronous, metachronous, or other) and site(s) of metastases and pre-treatment albumin (Guckenberger et al., 2020). Breast, prostate, and kidney were defined as favorable primary tumor sites as previously described and validated (Zucker et al., 2018). All six validated variables of the TEACHH model were also abstracted, including the number of hospitalizations within 3 months and the number of previous palliative chemotherapy courses (Krishnan et al., 2014). Radiation treatment variables captured included radiation modality (stereotactic radiation and/or intensity-modulated radiation) and radiation biologically effective dose averaged over all irradiated sites with an alpha/beta ratio of 10.

Following radiation, systemic therapy was administered at the discretion of the treating oncologist and was classified as none, non-cytotoxic (hormonal, targeted, and/or immunotherapy), or cytotoxic chemotherapy-containing regimens.

2.2 FaceAge 1.0

The FaceAge 1.0 deep learning pipeline was previously described (Bontempi et al., 2025). Briefly, this pipeline involves detecting the face using a cascaded convolutional neural network and then encoding the face into a feature vector to allow for age prediction using a regression algorithm. FaceAge 1.0 was trained using over 50,000 face photographs from the publicly available IMDb-Wiki dataset.

As part of the radiation planning process, a consumer-grade digital camera was used to obtain a frontal facial digital photograph of all 177 patients in the CT simulation suite under ambient lighting for patient identification. Ten photographs did not meet quality assurance standards described previously and were not included in the FaceAge analysis (Bontempi et al., 2025).

2.3 FaceAge 2.0

FaceAge 2.0 was developed from the FAHR-Face foundation model. The FAHR-Face model utilized a Vision Transformer architecture with masked autoencoders to enable the self-supervised pre-training of over 40 million unlabeled facial images from the WebFace42M dataset (Haugg et al., 2025; Zhu et al., 2021). This approach learns facial representations by randomly masking image patches and training the model to reconstruct them. The foundation model was fine-tuned for biological age

estimation on 8 public age-estimation datasets and further refined on the UTK and AgeDB datasets (Haugg et al., 2025; 2026).

Both FaceAge 1.0 and FaceAge 2.0 were downloaded from the publicly available GitHub repositories `https://github.com/AIM-Harvard/FaceAge` and `https://github.com/AIM-Harvard/FAHR-Face` and were executed without modification by the computer science collaborators on the study team. The computer scientist was blinded to calendar age and clinical outcomes at the time of prediction.

2.4 Statistical and Survival Analysis

The primary endpoint was overall survival from the date of radiation oncology consultation to date of death from any cause or last follow-up. Overall survival was estimated using the Kaplan–Meier method and compared using log-rank tests. Univariable Cox proportional hazards regression was performed for each covariate. FaceAge 1.0 and FaceAge 2.0 were analyzed as continuous variables, with hazard ratios reported per year (Bontempi et al., 2025). Racial bias was assessed using ANCOVA regression of FaceAge on chronological age with race as a covariate. Independent and paired t-tests compared FaceAge-chronological age differences between racial groups and between algorithm versions.

Harrell's concordance index (C-index) was calculated for each univariable model. Variables with $p < 0.10$ on univariable Cox regression were entered into Cox multivariable analyses. The incremental value of FaceAge beyond chronological age was evaluated by comparing the C-indices. Collinearity between FaceAge and chronological age was assessed by Pearson's correlation coefficient and Pearson's r-squared. Statistical analysis was done using Stata 13.1 and R version 4.3.2. Differences between C-indices were assessed with bootstrap 95% confidence intervals with 1000 resamples. Two-sided p values < 0.05 were considered statistically significant.

To determine the relationship between FaceAge 1.0, FaceAge 2.0, and chronological age with ECOG performance status, additional sensitivity analyses were performed using Spearman rank correlations. Post-radiation systemic therapy intensity was ordinally coded as 0=none, 1=non-cytotoxic therapy, and 2=chemotherapy and was correlated with baseline variables using Spearman correlations.

3 Results

Among 177 patients, the most common primary sites were lung ($n = 60$), prostate ($n = 28$), and breast ($n = 19$), with a median of 1 metastasis treated. Most patients were White ($n = 145$; 82%), 20 (11%) were Black, and 12 (7%) were recorded as other race. At a median follow-up of 44.0 months, median survival was 42.8 months (95% CI, 29.8 to 53.7) and 5-year overall survival was 39.6% (95% CI, 31.2% to 47.9%). The median chronological age was 70 years (IQR, 62–79), while the median FaceAge 1.0 was 73 years (IQR, 64–78). Overall survival by clinical and treatment characteristics is summarized in Table 1.

3.1 Evidence of Race-Associated Algorithmic Underestimation

During routine quality-assurance review of the output of FaceAge 1.0, patients were sorted by their estimated facial age, which revealed an apparent pattern of biological-age underestimation for Black patients. A striking visual pattern was observed and raised concern for systematic error and algorithmic bias, prompting formal testing. White patients had an average FaceAge 1.0 estimate 1.5 years older than chronological age ($p = 0.050$). The FaceAge 1.0 estimate was 6.8 years younger than chronological age in Black patients ($p = 0.006$). On multivariable ANCOVA regression adjusting for chronological age, Black race was independently associated with a 10.6-year lower FaceAge 1.0 score compared with White patients (95% CI, -14.2 to -7.1; $p < 0.001$) (Table 2). Based on these surprising findings, we consulted the FaceAge developers for assistance with troubleshooting our analysis pipeline. In the absence of identifying technical errors, we were advised to utilize FaceAge 2.0, which was trained on a larger and more racially inclusive dataset.

3.2 FaceAge 2.0 Attenuates the Race-Associated Difference

The median FaceAge 2.0 was 74 years (IQR, 64 to 81). After implementing the new algorithm, the mean FaceAge 2.0 estimate for Black patients was 1.4 years older than calendar age ($p = 0.62$) (Table 2). For Black patients, FaceAge 2.0 scored 8.2 years higher than FaceAge 1.0 ($p = 0.027$). In multivariable ANCOVA regression adjusting for chronological age, the residual coefficient for Black race was reduced from -10.6 years (FaceAge 1.0) to -3.0 years (95% CI -7.3 to +1.2; $p = 0.16$) with FaceAge 2.0.

3.3 External Validation of FaceAge 1.0 and FaceAge 2.0

FaceAge 1.0 (Pearson $r = 0.64$, $p < 0.001$) and FaceAge 2.0 (Pearson $r = 0.75$, $p < 0.001$) both showed a strong positive correlation with chronological age. Variance inflation factors (VIF) of FaceAge 1.0 and FaceAge 2.0 were 1.71 and 2.29, respectively, suggesting acceptable collinearity (VIF< 5) (Supplementary Figure S1a; Supplementary Figure S1b).

Chronological age (HR 1.04 per year, 95% CI 1.02–1.06; $p < 0.001$), FaceAge 1.0 (HR 1.03 per year, 95% CI 1.01–1.05; $p = 0.004$), and FaceAge 2.0 (HR 1.03 per year, 95% CI 1.01–1.04; $p = 0.002$) were all significant predictors of survival on univariable Cox regression (Table 3). For both model versions, younger FaceAge groups had longer overall survival than older FaceAge groups (Figure 1A-D; Table 1).

The C-indices of chronological age, FaceAge 1.0, and FaceAge 2.0 were 0.61, 0.59, and 0.60, respectively, when used as single variables. Other significant variables on univariable analysis included ECOG performance status (HR 1.69, 95% CI 1.32–2.16; $p < 0.001$), unfavorable primary tumor site (HR 2.52, 95% CI 1.51–4.21; $p < 0.001$), albumin (HR 0.48, 95% CI 0.35–0.67; $p < 0.001$), liver metastases (HR 2.35, 95% CI 1.28–4.32; $p = 0.006$), and post-radiation non-cytotoxic systemic therapy (HR 0.51, 95% CI 0.29–0.91; $p = 0.02$) (Table 3).

3.4 Relationship with Performance Status and Post-Radiation Therapy

FaceAge 2.0 was found to have a greater Spearman correlation with ECOG performance status ($\rho = 0.345$, $p < 0.0001$) than chronological age ($\rho = 0.283$, $p = 0.0001$), whereas FaceAge 1.0 had a Spearman correlation nearly identical to chronological age ($\rho = 0.283$, $p = 0.0002$) (Supplementary Table S3). Patients with ECOG 0–1 performance status had a mean FaceAge 2.0 of 71.9 years, while patients with ECOG 2–4 had a mean of 78.0 years ($p = 0.008$). Mean chronological age was 68.4 years for ECOG 0–1 and 72.6 years for ECOG 2–4 ($p = 0.03$).

Post-radiation systemic therapy intensity was associated with calendar age ($\rho = -0.167$, $p = 0.03$) and with ECOG performance status ($\rho = -0.320$, $p < 0.001$), while FaceAge 1.0 ($\rho = -0.037$, $p = 0.64$) and FaceAge 2.0 ($\rho = -0.130$, $p = 0.09$) were not statistically significant (Supplementary Table S2).

3.5 Multivariable Analysis

On multivariable analysis, substituting FaceAge for calendar age did not meaningfully change the conclusions of the multivariable analyses (Table 3). When adding FaceAge to calendar age, both FaceAge 1.0 (HR 1.01, $p = 0.47$) and FaceAge 2.0 (HR 1.01, $p = 0.36$) lost statistical significance. In the Cox models including chronological age and either FaceAge version, chronological age, ECOG performance status, primary tumor site, liver metastases, and albumin remained statistically significant (Supplementary Table S1). The C-index for the Cox model with calendar age was 0.74 and statistical discrimination minimally increased to 0.75 when either FaceAge 1.0 or FaceAge 2.0 was added.

4 Discussion

FaceAge is a deep learning-based facial health recognition algorithm that represents a potential use case for applying artificial intelligence to enhance clinical prognostication. As part of clinical workflow, patients undergoing radiation simulation routinely have a facial photograph taken for

identification purposes. It seems intuitive that a clinical scoring of patient fitness contains prognostic information to supplement routinely assessed biomarkers such as chronological age (Kao et al., 2022; Bontempi et al., 2025). FaceAge is a standardized, objective and validated metric. Attention mapping suggests that FaceAge models heavily weight periorbital regions, anterior cheeks, nasal folds and temporalis muscles (Haugg et al., 2025).

To our knowledge, this study represents the first external validation of FaceAge in an oligometastatic cohort. Our findings largely validate prior results from the Harvard group. Prior studies showed that cancer patients present with a FaceAge about 5 years older than chronological age (Bontempi et al., 2025; Lee et al., 2026). In this study, the average FaceAge 2.0 biological age was 3.7 years older than average chronological age. On univariable analysis, FaceAge 2.0 was strongly correlated with overall survival with hazard ratios similar to those reported by Bontempi and colleagues (Bontempi et al., 2025; Lee et al., 2026). As a single datapoint, FaceAge captures unique information separate from calendar age and performance status, with a C-index for overall survival of 0.60. In the original report from Harvard Medical School, FaceAge as a single variable achieved an AUC of 0.61 for predicting 3-year survival in the MAASTRO dataset (Bontempi et al., 2025).

Unlike previously published datasets, FaceAge was not found to be more predictive of overall survival than calendar age. Calendar age was more strongly prognostic in this oligometastatic cohort than in prior unselected metastatic cancer prognostic models such as NEAT where age was not a statistically significant variable (Zucker et al., 2018). Unlike unselected metastatic cohorts, a substantial proportion of oligometastatic patients achieve long-term disease control, making competing mortality from other causes a significant contributor to overall survival (Kao et al., 2023; 2024). For patients with oligometastases, calendar age, performance status, albumin and liver metastases have been shown to capture comorbidity burden, organ reserve and systemic inflammation that enhance survival predictions (Radigan et al., 2025). Both calendar age (Spearman rho -0.18, $p = 0.02$) and ECOG performance status (rho -0.30, $p < 0.001$) predicted post-radiation systemic therapy receipt while FaceAge 2.0 did not ($p = 0.09$), suggesting that community medical oncologists rely on traditional metrics when prescribing systemic treatment. Since the original FaceAge cohort focused on stage III lung cancer and the Harvard palliative cohort, it is also possible that FaceAge is less predictive in this 'in-between' oligometastatic population. As a single variable, albumin had a C-index of 0.66. Albumin is an indicator of hepatic reserve, immune function, and third-spacing and, despite its prognostic significance, is not widely captured in radiation oncology databases. Further research and validation of FaceAge 2.0 in a larger cohort of oligometastatic patients such as the ongoing OligoCare study is warranted (Alongi et al., 2024).

Although our cohort is relatively small, the percentage of Black patients (11%) is broadly representative of the population of the United States. Despite a small cohort of only 20 Black patients, we were surprised to find that FaceAge 1.0 systematically provided Black patients with younger biological age estimates (Table 2). A possible explanation is that FaceAge 1.0 was predominantly trained on the IMDb-Wiki dataset that was majority White and then was validated largely on data from the MAASTRO European biobank and Harvard thoracic and palliative cohorts (Bontempi et al., 2025). As artificial intelligence gains greater importance in medical care, avoiding bias is paramount to achieve equitable outcomes (Obermeyer et al., 2019). Image-based datasets are particularly vulnerable in the unequal representation of demographic characteristics and skin types (Wen et al., 2022).

FaceAge 2.0 substantially reduced racial disparities found with FaceAge 1.0 but retained the univariable association with survival (Table 2; Table 3). The newer FAHR-Face framework was pre-trained on a much larger facial-image corpus and optimized to learn more general facial representations prior to task-specific fine-tuning (Haugg et al., 2025). Given the small sample of Black patients in this study, further external fairness assessment in larger and more diverse cohorts is still needed.

More recent studies indicate that facial aging can also be more useful as a longitudinal signal rather than as a single time-point measure. Haugg et al. (2026) estimated Face Aging Rate from serial clinical photographs in 2,276 cancer patients and reported that faster increases in FaceAge were linked to poor overall survival, after controlling for clinical factors. No longitudinal FaceAge trajectories could be assessed in our study because only one baseline simulation photograph was used. This is a natural direction for future work in patients with oligometastatic disease, who tend to follow up for several years. Facial biomarkers could be routinely incorporated into multimodal prognostic models that combine structured and unstructured clinical, treatment, demographic, and genomic information automatically extracted from routine records (Jee et al., 2024).

This study has several limitations. Precision of subgroup analysis is limited by the modest sample size of 177 patients, including only 20 Black patients. The single-institution design with photographs primarily taken by a designated simulation therapist limits generalizability. Photographs were taken with a consumer-grade digital camera under ambient lighting in the simulation suite rather than standardized conditions. The retrospective nature of this analysis precludes causal inference between FaceAge and treatment decisions.

In conclusion, FaceAge was externally validated as a univariable prognostic biomarker in oligometastatic cancer. FaceAge 2.0 substantially reduced racial bias observed with FaceAge 1.0. Chronological age was highly prognostic in this small but relatively long-surviving cohort, which may explain why FaceAge did not contribute independent multivariable discrimination beyond established prognostic factors. These findings support continued development and validation of FaceAge 2.0 as a more equitable facial-aging biomarker for oncology studies.

Ethics Statement

This retrospective study was approved by the institutional review board at the participating institution; facial photographs and linked clinical data were obtained during routine care, were not publicly released because the photographs are identifiable, and race was taken from the electronic health record rather than inferred from the images.

Reproducibility Statement

Cohort definitions, image-quality criteria, model versions, statistical methods, and software are described in the Methods and supplement; identifiable facial photographs and linked clinical records cannot be publicly released, so reproduction requires an appropriately approved dataset with comparable inputs.

AI Use Statement

Generative AI tools were used to assist with manuscript organization, editing, reference formatting, and conversion to the required submission format; no generative AI was used to generate patient data or the reported statistical analyses.

References

Filippo Alongi, Luca Nicosia, Umberto Ricardi, et al. Acute toxicity in patients with oligometastatic cancer following metastasis-directed stereotactic body radiotherapy: An interim analysis of the e2-radiate oligocare cohort. *Radiotherapy and Oncology*, 199:110466, 2024. doi: 10.1016/j.radonc.2024.110466.

Evgeny Bobrov, Anastasia Georgievskaya, Kirill Kiselev, et al. Photoageclock: deep learning algorithms for development of non-invasive visual biomarkers of aging. *Aging*, 10(11):3249–3259, 2018. doi: 10.18632/aging.101629.

Davide Bontempi, Oszkar Zalay, Danielle S. Bitterman, et al. Faceage, a deep learning system to estimate biological age from face photographs to improve prognostication: a model development and validation study. *The Lancet Digital Health*, 7(6):100870, 2025. doi: 10.1016/j.landig.2025.03.002.

William Dale, Heidi D. Klepin, Grant R. Williams, et al. Practical assessment and management of vulnerabilities in older patients receiving systemic cancer therapy: Asco guideline update. *Journal of Clinical Oncology*, 41(26):4293–4312, 2023. doi: 10.1200/JCO.23.00933.

Matthias Guckenberger, Yolande Lievens, Anne B. Bouma, et al. Characterisation and classification of oligometastatic disease: a european society for radiotherapy and oncology and european organisation for research and treatment of cancer consensus recommendation. *The Lancet Oncology*, 21(1):e18–e28, 2020. doi: 10.1016/S1470-2045(19)30718-1.

Fridolin Haugg, Grace Lee, Jing He, et al. Foundation artificial intelligence models for health recognition using face photographs (fahr-face), 2025.

Fridolin Haugg, Grace Lee, Jing He, et al. Face aging rate quantifies change in biological age to predict cancer outcomes. *Nature Communications*, 17(1):3487, 2026. doi: 10.1038/s41467-025-66758-w.

Justin Jee, C. Fong, K. Pichotta, et al. Automated real-world data integration improves cancer outcome prediction. *Nature*, 636(8043):728–736, 2024. doi: 10.1038/s41586-024-08167-5.

Johnny Kao, Alyson Zucker, Michael Urso, et al. Improving survival prognostication in patients with metastatic cancer through clinical judgment. *Anticancer Research*, 42(3):1397–1401, 2022. doi: 10.21873/anticanres.15609.

Johnny Kao, Michael Sahagian, Vishal Gupta, Symeon Missios, and Aman Sangal. Long-term disease-free survival following comprehensive involved site radiotherapy for oligometastases. *Frontiers in Oncology*, 13:1267626, 2023. doi: 10.3389/fonc.2023.1267626.

Johnny Kao, Paul Eckardt, John McEachron, Christina Atalla, and Aman Sangal. Predicting long-term survival following involved site radiotherapy for oligometastases. *Oncology Letters*, 27(2):82, 2024. doi: 10.3892/ol.2024.14216.

Meera S. Krishnan, Zachary Epstein-Peterson, Yuh-Hwa Chen, et al. Predicting life expectancy in patients with metastatic cancer receiving palliative radiotherapy: the teachh model. *Cancer*, 120 (1):134–141, 2014. doi: 10.1002/cncr.28408.

Grace Lee, Fridolin Haugg, Davide Bontempi, et al. Faceage as a biomarker for prognosis and treatment stratification in a large-scale oncology cohort. *Journal of the National Cancer Institute*, 118(5):811–821, 2026. doi: 10.1093/jnci/djaf323.

Allison Magnuson, Kah Poh Loh, Fiona Stauffer, et al. Geriatric assessment for the practicing clinician: The why, what, and how. *CA: A Cancer Journal for Clinicians*, 74(6):496–518, 2024. doi: 10.3322/caac.21864.

Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019. doi: 10.1126/science.aax2342.

David A. Palma, Alexander V. Louie, and George B. Rodrigues. New strategies in stereotactic radiotherapy for oligometastases. *Clinical Cancer Research*, 21(23):5198–5204, 2015. doi: 10.1158/1078-0432.CCR-15-0822.

Ryan Radigan, C. S. Kao, Michael Krainock, et al. Long-term survival and undetectable circulating tumor dna following comprehensive involved site radiotherapy for oligometastases. *Scientific Reports*, 15(1):6126, 2025. doi: 10.1038/s41598-025-88266-z.

David Wen, Saad M. Khan, Anthony J. Xu, et al. Characteristics of publicly available skin cancer image datasets: a systematic review. *The Lancet Digital Health*, 4(1):e64–e74, 2022. doi: 10.1016/S2589-7500(21)00252-1.

Zheng Zhu, Guan Huang, Jiankang Deng, et al. Webface260m: A benchmark unveiling the power of million-scale deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10492–10502, 2021. doi: 10.1109/CVPR46437.2021.01035.

Alyson Zucker, C. J. Tsai, Joseph Loscalzo, Pierre Calves, and Johnny Kao. The neat predictive model for survival in patients with advanced cancer. *Cancer Research and Treatment*, 50(4): 1433–1443, 2018. doi: 10.4143/crt.2017.223.

APPENDIX / SUPPLEMENTARY MATERIAL

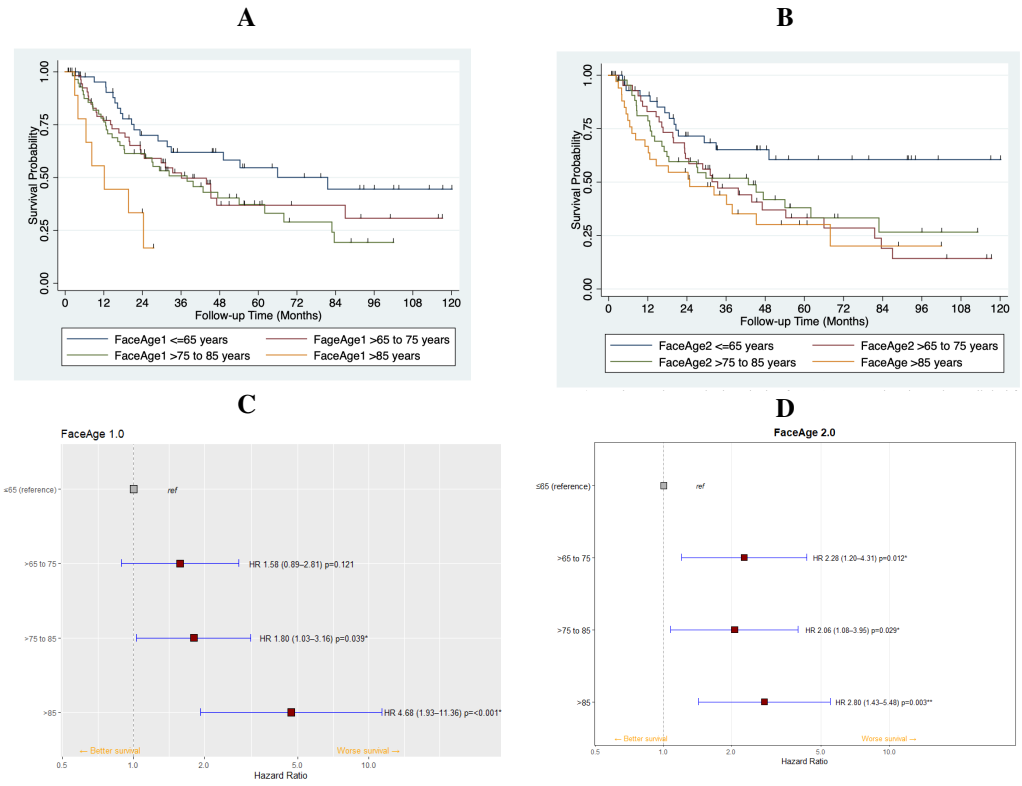

Figure 1: Prognostic Performance of FaceAge 1.0 and FaceAge 2.0 in Patients with Oligometastatic Cancer. A) Kaplan-Meier survival analysis of FaceAge 1.0 estimation using a digital face photograph. B) Kaplan-Meier survival analysis of FaceAge 2.0 estimation using a digital face photograph. C) Forest plots of FaceAge 1.0 risk groups on overall survival. D) Forest plots of FaceAge 2.0 risk groups on overall survival.

Table 1: Kaplan-Meier Estimates of Overall Survival in Patients with Oligometastatic Cancer, Stratified by Clinical and Treatment Characteristics. Median survival times (in months) and 5-year survival percentages are shown. P-values represent log-rank tests comparing survival curves between groups.

	Number	Percent (%)	median overall survival (months)	5-year overall survival	P value
Total	177	100	42.8	39.6%	
ECOG Performance Status					$< 0.01^*$
0–1	129	73	45.2	42.2%	
2–4	48	28	29.8	34.8%	
Gender					0.13
Male	98	55	49.2	45.3%	
Female	79	45	28.9	32.1%	
Race					
White (ref)	145	82	36.1	36.7%	
Black	20	11	45.2	49.1%	0.33
Other	12	7	70.9	53.5%	0.34
Age					$< 0.01^*$
< 70	85	48	48.8	44.6%	
≥ 70	92	52	37.9	35.0%	
FaceAge 1.0					$< 0.01^*$
< 70	69	39	66.0	50.3%	
≥ 70	98	55	31.2	33.8%	
Missing	10	6	43.2	21.4%	
FaceAge 2.0					$< 0.01^*$
< 70	65	37	81.5	52.1%	
≥ 70	102	58	32.3	33.7%	
Missing	10	6	43.2	21.4%	
Primary site					< 0.01
Lung	60	34	28.9	34.7%	0.51
Prostate	28	16	81.5	64.0%	< 0.01
Breast	19	11	82.9	55.7%	0.06
Other	70	40	24.6	28.2%	ref
Favorable primary tumor					< 0.01
Breast, Prostate or Kidney	50	28	81.5	59.2%	
Other	127	72	24.6	31.3%	
Metastasis site					
Bone	61	34	43.9	41.4%	0.29
Brain	47	27	24.8	29.6%	0.27
Distant Lymph Nodes	36	20	49.2	47.3%	0.27
Lung	34	19	42.8	35.8%	0.87
Liver	16	9	10.4	29.2%	< 0.01
Prior Palliative Chemotherapy Regimens					0.20
0–1	146	82	45.2	40.9%	
≥ 2	31	18	29.8	32.9%	
Hospitalized During Prior 3 Months					0.51
No	114	64	43.9	40.7%	
Yes	63	36	31.7	37.2%	
Baseline Albumin					$< 0.01^*$
≥ 3.8 g/dl	78	44	45.3	50.3%	
< 3.8 g/dl	84	47	14.9	28.3%	
Missing	15	8	49.2	38.6%	
EORTC Oligometastasis Classification					0.35
Synchronous	75	42	53.8	43.2%	Ref
Metachronous	63	36	32.1	35.9%	0.23
Other (Repeat or Induced)	39	22	31.7	36.0%	0.21
Number of Metastases Treated					0.20
0 to 3	161	90	37.9	38.6%	
4 to 5	16	9	47.4	45.7%	
RT dose (BED10)					0.27
< 75 Gy	118	67	45.2	43.9%	
≥ 75 Gy	59	33	28.9	29.6%	
Post-RT systemic therapy					0.01
None (ref)	38	21	18.6	39.4%	
Non-cytotoxic systemic therapy (hormonal, immunotherapy and/or targeted)	73	41	67.9	55.3%	0.02
Chemotherapy (with or without non-cytotoxic systemic therapy)	66	37	31.2	25.6%	0.94

* Analyzed as a continuous variable.

Table 2: Age-Adjusted Differences in FaceAge 1.0 and FaceAge 2.0 Predictions by Race.

	White ($n = 138$)	**Black** ($n = 20$)	**Other** ($n = 9$)
Chronological age	71.0 ± 11.4	65.8 ± 7.7	61.9 ± 11.2
FaceAge 1.0			
Mean ± SD	72.5 ± 9.5	59.0 ± 9.5	63.3 ± 10.3
Difference from chronological age	+1.5	-6.8	+1.4
One-sample t-test p	0.050	0.006	0.54
ANCOVA coefficient vs White	ref	-10.6 (-14.2 to -7.1)	-4.2 (-9.3 to 0.9)
ANCOVA p-value	—	< 0.001	0.11
FaceAge 2.0			
Mean ± SD	74.8 ± 13.2	67.2 ± 12.4	69.6 ± 15.2
Difference from chronological age	+3.8	+1.4	+7.7
One-sample t-test p	< 0.001	0.62	0.003
ANCOVA coefficient vs White	ref	-3.0 (-7.3 to 1.2)	+2.9 (-3.2 to 9.0)
ANCOVA p-value	—	0.16	0.35
FA 2.0 vs FA 1.0 (paired comparison)			
Mean increase (FA2.0 - FA1.0)	+2.3	+8.2	+6.3
Paired t-test p	0.008	0.027	0.056

Age-adjusted ANCOVA with chronological age as covariate. Values are mean ± SD unless otherwise noted. **Overall race effect:** FaceAge 1.0, $p < 0.001$; FaceAge 2.0, $p = 0.20$. ANCOVA = analysis of covariance. Differences from chronological age were calculated as FaceAge minus chronological age. ANCOVA coefficients represent the adjusted difference in FaceAge relative to White patients after accounting for chronological age.

Table 3: Univariable and Multivariable Cox Regression Analyses of Overall Survival in Patients with Oligometastatic Cancer. Variables with $p < 0.10$ on univariable analysis were considered for multivariable modeling. Calendar age, FaceAge 1.0, and FaceAge 2.0 were evaluated in separate multivariable models with the same clinical covariates.

Variable	HR	95% CI	P value	HR	95% CI	P value	HR	95% CI	P value	HR	95% CI	P value
FaceAge1.0 (continuous, per year)	1.03	1.01–1.05	0.004				1.04	1.01–1.06	0.005			
FaceAge2.0 (continuous, per year)	1.03	1.01–1.04	0.002							1.03	1.01–1.05	0.003
Age (continuous, per year)	1.04	1.02–1.06	< 0.001	1.03	1.01–1.05	0.003						
ECOG (continuous)	1.69	1.32–2.16	< 0.001	1.51	1.12–2.03	0.007	1.50	1.07–2.10	0.018	1.48	1.06–2.08	0.023
Primary site (breast, prostate or kidney vs. other)	2.52	1.51–4.21	< 0.001	2.37	1.35–4.15	0.003	2.26	1.17–4.38	0.015	2.01	1.04–3.86	0.036
Liver metastases (no vs. yes)	2.35	1.28–4.32	0.006	2.11	1.13–3.95	0.02	1.98	1.01–3.85	0.046	2.41	1.24–4.68	0.010
Albumin (continuous, per g/dl)	0.48	0.35–0.67	< 0.001	0.69	0.49–0.98	0.04	0.61	0.42–0.88	0.009	0.61	0.42–0.88	0.009
Post-RT systemic therapy												
None (ref)												
Non-cytotoxic systemic therapy (hormonal, immunotherapy and/or targeted)	0.51	0.29–0.91	0.02	0.68	0.36–1.27	0.22	0.60	0.32–1.14	0.12	0.62	0.33–1.18	0.14
Chemotherapy (with or without non-cytotoxic systemic therapy)	0.98	0.58–1.68	0.95	1.22	0.66–2.29	0.53	0.91	0.49–1.71	0.78	1.12	0.59–2.09	0.74

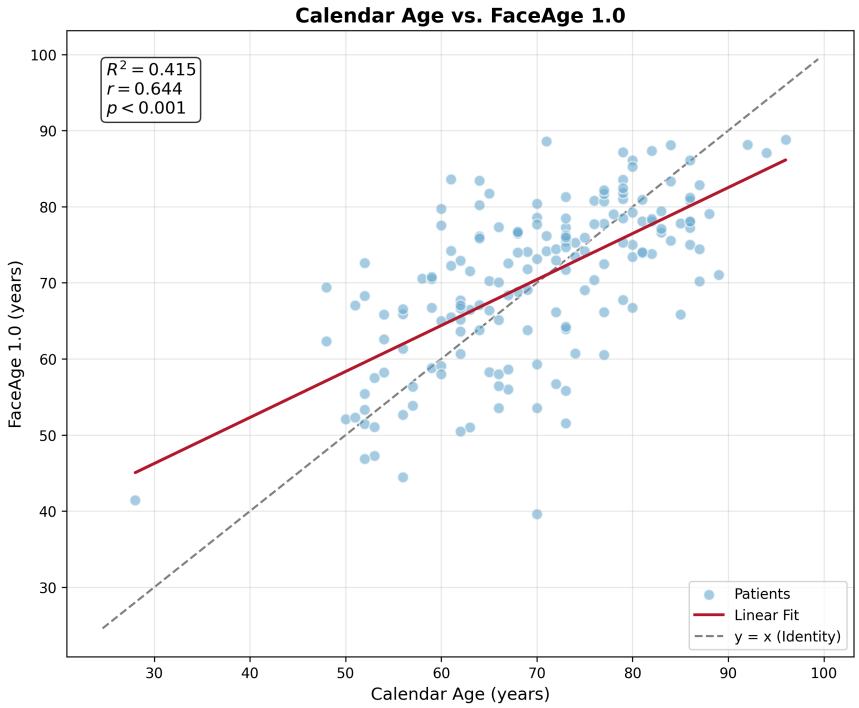

Figure S1a: Scatter plot of FaceAge 1.0 vs. Calendar Age

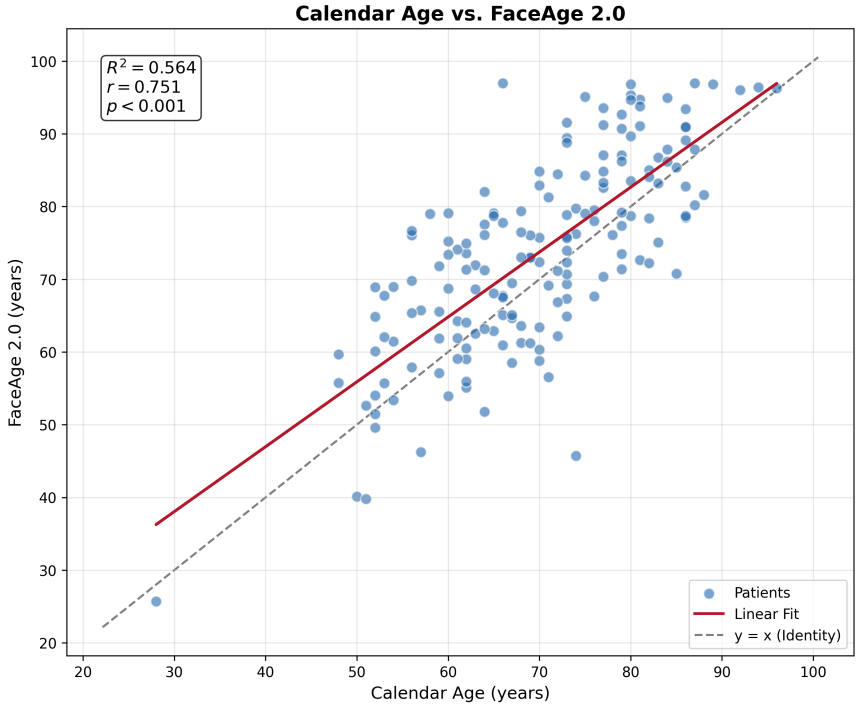

Figure S1b: Scatter plot of FaceAge 2.0 vs. Calendar Age

Table S1: Multivariable Cox models of overall survival including calendar age with either FaceAge 1.0 or FaceAge 2.0.

Variable	HR	95% CI	P value	HR	95% CI	P value
FaceAge1.0 (continuous, per year)	1.01	0.98–1.04	0.47			
FaceAge2.0 (continuous, per year)				1.01	0.99–1.04	0.36
Age (continuous, per year)	1.04	1.01–1.07	0.010	1.04	1.01–1.07	0.016
ECOG (continuous)	1.51	1.06–2.15	0.024	1.49	1.04–2.12	0.028
Primary site (breast, prostate or kidney vs. other)	2.07	1.07–3.99	0.030	1.99	1.04–3.80	0.038
Liver metastases (no vs. yes)	2.31	1.16–4.59	0.017	2.46	1.26–4.78	0.008
Albumin (continuous, per g/dl)	0.68	0.46–0.99	0.046	0.67	0.46–0.98	0.040
Post-RT systemic therapy						
None (ref)						
Non-cytotoxic systemic therapy (hormonal, immunotherapy and/or targeted)	0.63	0.33–1.19	0.16	0.63	0.33–1.19	0.15
Chemotherapy (with or without non-cytotoxic systemic therapy)	1.21	0.61–2.36	0.57	1.26	0.66–2.40	0.48

Table S2: Associations of calendar age, ECOG performance status, and FaceAge with post-radiation systemic therapy intensity.

A. Spearman Correlations with Post-Radiation Systemic Therapy Receipt

Variable	N	Spearman ρ	p-value
Calendar age	167	-0.167	0.03
ECOG	167	-0.320	< 0.001
FaceAge 1.0	167	-0.037	0.64
FaceAge 2.0	167	-0.130	0.09

B. Baseline Characteristics by Post-Radiation Systemic Therapy Group

Post-RT Systemic Therapy	N	Mean Age (SD)	Mean ECOG (SD)	Mean FA 1.0 (SD)	Mean FA 2.0 (SD)
None	38	71.7 (13.4)	1.66 (0.94)	70.6 (11.1)	75.4 (15.4)
Non-cytotoxic	67	71.2 (10.9)	1.06 (0.80)	70.5 (10.8)	74.6 (13.1)
Chemotherapy	62	67.2 (9.8)	0.85 (0.70)	70.0 (10.1)	71.4 (12.2)

Spearman correlations between baseline variables and post-radiation systemic therapy intensity (0=none, 1=non-cytotoxic, 2=chemotherapy). $N = 167$ patients with complete data.

Table S3: Spearman Correlations of FaceAge and Calendar Age with ECOG Performance Status.

Comparison	N	ρ	p
FaceAge 1.0 vs. ECOG	167	0.283	0.0002
FaceAge 2.0 vs. ECOG	167	0.345	< 0.0001
Calendar age vs. ECOG	177	0.283	0.0001