Nuclear morphology is a deep learning biomarker of senescence across tissues and species

Indra Heckenbach  
University of Copenhagen

Garik Mkrtchyan  
University of Copenhagen

Michael Ben Ezra  
University of Copenhagen

Daniela Bakula  
University of Copenhagen

Jakob Madsen  
University of Copenhagen  https://orcid.org/0000-0002-2841-7284

Malte Nielsen  
Gubra

Denise Oró  
Gubra

M Laura Idda  
National Institute on Aging

Myriam Gorospe  
Nat Inst on Aging, NIH

Laust Mortensen  
University of Copenhagen

Eric Verdin  
Buck Institute for Research on Aging

Rudi Westendorp  
University of Copenhagen

Morten Scheibye-Knudsen  
mscheibye@sund.ku.dk  
University of Copenhagen  https://orcid.org/0000-0002-6637-1280

Article

Keywords: cellular senescence, nuclear morphology, age-related diseases

Posted Date: November 11th, 2021
Nuclear morphology is a deep learning biomarker of senescence across tissues and species

Indra Heckenbach\textsuperscript{1,2,3}, Garik V Mkrtchyan\textsuperscript{1}, Michael Ben Ezra\textsuperscript{1,4}, Daniela Bakula\textsuperscript{1}, Jakob Sture Madsen\textsuperscript{1}, Malte Hasle Nielsen\textsuperscript{5}, Denise Oró\textsuperscript{5}, M Laura Idda\textsuperscript{6,7}, Myriam Gorospe\textsuperscript{6}, Laust Mortensen\textsuperscript{4,8}, Eric Verdin\textsuperscript{2}, Rudi Westendorp\textsuperscript{4,8}, Morten Scheibye-Knudsen\textsuperscript{1,3*}

\textsuperscript{1}Center for Healthy Aging, Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark
\textsuperscript{2}Buck Institute for Research on Aging, Novato, CA, USA
\textsuperscript{3}Tracked.bio, Copenhagen, Denmark
\textsuperscript{4}Methods and Analysis, Statistics Denmark, Copenhagen, Denmark
\textsuperscript{5}Gubra, Hørsholm, Denmark
\textsuperscript{6}Laboratory of Genetics and Genomics, National Institute on Aging Intramural Research Program, National Institutes of Health, Baltimore, MD, USA
\textsuperscript{7}Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Sassari, Italy
\textsuperscript{8}Department of Public Health, University of Copenhagen, Copenhagen, Denmark

*Correspondence: mscheibye@sund.ku.dk
Cellular senescence is a critical component of aging and many age-related diseases, but understanding its role in human health is challenging in part due to the lack of exclusive or universal markers. Using neural networks, we achieve high accuracy in predicting senescence state and type from the nuclear morphology of DAPI-stained human fibroblasts, murine astrocytes, murine neurons, and fibroblasts derived from premature aging diseases in culture. After generalizing this approach, the predictor recognizes an increasing rate of senescent cells with age in H&E-stained murine liver tissue and human dermal biopsies, suggesting that alterations in nuclear morphology is a universal feature of senescence. Evaluating corresponding medical records reveals that individuals with a higher rate of senescent cells have a significantly decreased rate of malignant neoplasms, lending support for the protective role of senescence in limiting cancer development. Additionally, we find a positive association with lower significance for other conditions, including osteoporosis, osteoarthritis, hypertension, cerebral infarction, hyperlipidemia, and hypercholesteremia. In sum, we introduce a predictor of cellular senescence based on nuclear morphology that is applicable across tissues and species and is associated with health outcomes in humans.
**Introduction**

Cellular senescence is widely recognized as a fundamental process in aging, both as a primary causal factor in the decline of tissue homeostasis and as a consequence of other aging processes such as inflammation and DNA damage\(^1-^3\). Due to its critical role in disease etiology, senescence is increasingly recognized as a target for pharmaceutical intervention\(^4\). It also serves as a biomarker for aging\(^5\), possibly providing a more nuanced measure of age-related health in model organisms beyond simple chronological age. However, the role of senescence in human health is not clearly understood. Senescent cells present a complex and diverse phenotype, which varies significantly by cell type and source\(^6,^7\). There is considerable overlap between molecular factors that associate with senescence, DNA damage, inflammation, and other processes\(^8,^9\). Some of the most common markers of senescence are senescence-associated β-galactosidase activity (SA-β-gal) and increased levels of the cell cycle inhibitors p16 and p21. Nevertheless, there is no single marker that reliably and consistently identifies senescence\(^10-^12\). Importantly, senescent cells often exhibit an altered morphology, including expanded nuclei and an irregular, flattened appearance\(^13,^14\), making senescence amenable to analysis with computer vision and machine learning methods\(^15\).

We present deep learning models that can predict cellular senescence with high accuracy based on nuclear morphology. These methods can further distinguish between multiple types of senescence, including radiation-induced damage and replicative exhaustion. Notably, predicted senescence correlates substantially with SA-β-gal, p16, p21, p53, and DNA damage markers γH2AX and 53BP1 foci counts. Our senescence predictor was developed using normal human fibroblast lines, but it also identifies increased senescence when applied to multiple types of premature aging diseases, including Hutchinson-Gilford progeria syndrome, ataxia...
telangiectasia, and Cockayne syndrome. We also evaluated the predictor on mouse astrocytes and neurons and found it indicates increased senescence in cells subjected to ionizing radiation, confirming its relevance to different cell types and organisms. These methods were further applied to H&E-stained mouse liver tissue, where we found an increasing rate of senescence with age. In addition, H&E-stained human tissue sections showed an age-dependent increase in senescence. Using the Danish National Patient Register, which records all ambulatory and in-patient contacts with Danish hospitals, we investigated how predicted senescence relates to human disease. In our study of 169 individuals, we found a highly statistically significant relationship between malignant neoplasm incidence and fewer predicted senescent cells, which fits the hypothesis that senescence is a mechanism to limit cancer. While oncogenic events are associated with the formation of senescent cells, we speculate that individuals with higher propensity toward developing senescent cells have reduced formation of malignant neoplasm and are at lower risk of cancer. We also found a weaker association between predicted senescence and other conditions, including osteoporosis, osteoarthritis, hypertension, cerebral infarction, hyperlipidemia, and hypercholesteremia.

Results

Three dermal fibroblast cell lines were induced to senescence by ionizing radiation (IR) or passaged until they reached replicative senescence (RS) (Fig. S1a, b, c). To confirm that the IR treated cells were senescent we evaluated levels of senescence markers p16, p21, p53 and interleukin-6 by immunohistochemistry and/or qPCR and found that IR led to a significant increase in these markers (Fig. S1d, e, f, g, h, i, j). Importantly, IR induced growth arrest as measured by cell counts for 1 week after IR treatment (Fig. S1k). Furthermore, DAPI intensity has been shown to decrease with senescence and this was indeed the case in RS and IR.
treated cells (Fig. S11). Using these established senescent models, DAPI-stained nuclei from IR and RS senescent cells were imaged with a high-content microscope. Nuclei were detected using a deep convolutional neural network based on U-net, which produced output images containing the detected nuclear regions. Each detected nucleus was extracted for subsequent analysis. We applied several methods to normalize features in images, such as removing the background, standardizing the size of the nuclei, and even masking inner details of the nuclei (Fig. 1a, b).

**Senescent Cells Display Altered Nuclear Morphology**

A morphological analysis of the detected nuclei was performed to compare control cells to those that were senescent. There was a significant difference in nuclear area for each of the three groups as previously reported \(^{13}\). In addition, IR senescent cells were significantly larger than RS cells (Fig. 1c). Aging and certain premature aging diseases have been associated with greater irregularities or folds in the nuclear envelope \(^{20,21}\). We therefore evaluated convexity, which is a ratio of nuclear perimeter to convex hull perimeter, as a measure of the nuclear envelope regularity. Convexity showed the shape of control cells were more regular compared to both IR and RS, which had a more irregular boundary (Fig. 1d). RS had the lowest convexity value, indicating the highest irregularity. This suggests convexity is another measure of senescence, with lower values corresponding to increased senescence. In addition, we looked at aspect ratio, a measure of width compared to height (measured as the longest compared to shortest dimensions of a minimized, rotated rectangle around each nucleus) and found that both IR and RS had higher values compared to control (Fig. 1e). We evaluated area and convexity per nuclei, observing overlapping clusters for the three states with area of RS overlapping both control and IR, and convexity of RS and IR overlapping (Fig. 1f). Interestingly, the distribution of
Figure 1 Nuclear morphology is an accurate senescence predictor in cultured cells. a Analysis workflow. b Sample nuclei for controls, replicative senescence (RS) and ionizing radiation (IR) induced senescent cells. c Area of identified nuclei (n=6,976-68,971, mean ± 95% CI). d Convexity of identified nuclei (n=6,976-68,971, mean ± 95% CI). e Aspect ratio of identified nuclei (n=6,976-68,971, mean ± 95% CI). f Scatter plot of individual nuclei, with overall distributions for each at the top and right margins. g Cell cycle analysis after exposure to several doses of IR; mn: multinucleated cells. h Accuracy of a deep neural network (DNN) predictor on validation data. i Receiver operating characteristics (ROC) curve of the DNN. j Percent of nuclei in each state classified as senescent for independent cell lines. k Distribution of prediction probabilities for several doses of IR for three fibroblast cell lines (n=38,284-106,132). l Distribution of p21 intensities for several doses of IR for three fibroblast cell lines (n=38,284-106,132). m Distribution of PCNA intensities for several doses of IR for three fibroblast cell lines (n=38,284-106,132). n Correlation between predicted senescence and nearby SA-β-gal regions, showing all and 90% confidence predictions only for RS and IR groups. o Correlation between predicted senescence and multiple markers, showing all, filtered for markers with strong signals, and filtered with 90% confidence predictions only. p Accuracy of DNNs trained and predicting after different normalization methods. q Correlation between morphological metrics and predicted senescence by class, BG: background.
the area of the IR senescent cells was bimodal, with the lower mode matching RS and a higher mode at almost twice the area of the RS, perhaps suggesting IR induced aneuploidy or stalling at the G2 checkpoint of the cell cycle (**Fig. 1f**, upper histogram distribution of joint scatter plot).

To further explore this hypothesis, we induced senescence with multiple IR doses and utilized flow cytometry to study the cell cycle. Remarkably, we observed a dose-dependent increase in G2 and corresponding loss of G1 and S-phase cells 10 days after the IR treatment (**Fig. 1g**), indicating that IR induction leads to G2 stalled senescent cells as previously suggested. Simple nuclear morphological measures appear to be a viable method for assessing cellular senescence in culture.

**Deep Learning Classifiers Accurately Predicts Senescence Based on DAPI staining**

Given the rich structure of nuclei and potentially broad set of features, we applied deep neural networks to better assess senescence. A custom convolutional neural network was trained using 80% of the samples and 20% was held out for validation. After seeing accuracy converge to a steady level, the model was applied to validation data. We also compared our custom network to Xception, one of the top performing models for image classification that has been often applied to biomedical classification. Xception achieved superior results with an f1-score of 94%, accuracy of 95%, and AUC of 0.99 with validation data (**Fig. 1h, i**). To eliminate any potential overfitting on the experimental context and cell lines, we evaluated the model on an independent data set of two additional cell lines, which were prepared and imaged separately. This achieved an f1-score of 92%, accuracy of 94%, and AUC of 0.96 (**Fig. S2a, b, c**). The predictor identifies senescence for 12.7% of control, 92.0% of RS, and 95.6% for IR (**Fig. 1j**).
To better characterize the performance with senescent phenotypes induced by multiple levels of stress, we applied the DNN predictor to cells exposed to different doses of radiation. All levels were predicted to be senescent, and there was a 9.7% mean increase between 5 and 10 Gy, but 10 Gy to 20 Gy show a similar prediction score (Fig. 1k). PCNA declines with increasing dose and p21 increases (Fig. 1l, m). Predicted senescence and the two markers align with experimental conditions, but the predictor appears to track p21 expression more closely than PCNA. This experiment indicates that the treatment dose influences the senescent phenotype up to 10 Gy, a dose commonly used for senescence induction.

In another experiment, a deep neural network was trained to detect control compared to different senescent types, IR and RS. Xception, trained like the dual state model above, produced a mean class accuracy of 78.6% in detection of the three states, with 83.3% for controls, 75.7% for RS, and 76.8% for IR (Fig. S2d, e). It achieved a relatively high AUC of 0.9 for RS and 0.95 for IR. In sum, nuclear morphology represents a strong predictor of both replicative and DNA damage induced senescence.
Deep Predictor Is Confirmed by Senescent Markers

To confirm the accuracy of the DNN, we evaluated the correlation between predictions and traditional markers of senescence, including SA-β-gal, p16, p21, and p53. Training a deep neural network to recognize SA-β-gal regions, we found SA-β-gal near nuclei for 64.1% of IR and 65.8% of RS compared to 19.6% for control, which roughly matches published rates for RS and controls. A correlation analysis revealed a Pearson coefficient of 0.39 for IR and 0.31 for RS between predicted senescence and SA-β-gal detected nearby, but when restricting to the treated cells with nearby SA-β-gal and controls without it, the correlation rose to 0.83 for IR and 0.67 for RS (Fig. 1n). Applying a 90% confidence filter (see sections below on confidence and deep ensemble methods), correlation rose to 0.96 for IR and 0.90 for RS, indicating the predictor is highly effective at recognizing senescence with detected SA-β-gal. On a per cell basis, we found moderate correlation between p16, p21, and p53 stain intensities and predicted senescence, but after applying a threshold to classify as positive markers and filtering out nuclei near the threshold (due to the broad overlap in the distribution of intensities, Fig. S1g, h, i), the correlation rose significantly to 0.69 for p16, 0.59 for p21, and 0.63 for p53 (Fig. 1o). We also applied confidence filtering, restricting nuclei to those with high predictive confidence, and found correlation of 0.86 for p16, 0.78 for p21, and 0.79 for p53, indicating high confidence predictions are effectively identifying senescent nuclei with p16, p21, and p53. The deep predictor is inferring senescence in agreement with multiple markers of senescence.
Nuclear Shape Is a Central Predictive Feature in Senescent Cells

Nuclei images contain several features that could be used for classification; however, it is unclear what the deep neural network is using as its basis for assessment. Nuclear area, staining intensity and even the image background itself could contain a signal that the neural network is picking up on. To provide some insight into how much these potential factors contribute to senescence classification, we trained several models based on reduced forms of the cutout library. Our base model already includes brightness standardization. First, the background of the nuclei was masked, by excluding all areas outside of the U-Net detected nuclear region. Next, we applied size normalization, such that the greater of the width and height was set to a standard pixel size. Finally, we converted the interior of nuclei to a single-color value, essentially masking all internal structure. With each reduction, we observed a slight decrease in classification accuracy when applied to independent test lines (Fig. 1p). The background masking produced 86% for the f1-score and 88% for accuracy, a small reduction indicating limited reliance of the background. With background masked and size normalized, a trained model produced 87% for f1-score and 88% for accuracy, showing area and size played little role in senescent detection. This model was further reduced by completely masking the internal structure of the nuclei, which led to an f1-score of 80% and accuracy of 78% (Fig. S2f, g). While masking was a significant reduction in accuracy, it is remarkable that so much information could be removed from nuclear images and still obtain a relatively accurate classification of senescence. These experiments suggest that classification is largely based on the overall shape of the nuclei. We explored this further by evaluating Pearson correlation between predicted senescence and several morphological metrics, finding that area was moderately correlated (despite being standardized by the predictor) but convexity and aspect ratio were weaker (Fig. 1q). The deep learning model appears to be picking up on the nuclear shape in a more sophisticated manner than simple morphometrics.
The final reduced model yields an overall accuracy of 78%, and it shows an imbalanced per class accuracy of 73.9% for control, 69.3% for RS, and 91.4% for IR. It maintains a good AUC of 0.88. With similar reductions, the three-state senescent type detector model shows overall accuracy of only 58% (Fig. S2h, i): 87.7% for controls, 56.1% for IR, but only 31.3% for RS. The AUC has declined to 0.71 for RS and 0.6 for IR. Despite lowering accuracy, the feature standardization and reduction makes the model less influenced by a large number of technical variations such as image intensity, choice of nuclear staining method, magnification and others that could impact the utility of the predictor.

Classification with Confidence

While overall accuracy per-nuclei was relatively high, a sizable number of nuclear images were ambiguous, which can be interpreted as the model being uncertain in its prediction. Extending neural networks with Bayesian properties has several advantages, most notably providing a measure of confidence for predictions. The Bayesian Neural Networks (BNN) allows for the construction of a posterior probability distribution which can be used for interval estimation, compared to a single prediction from a classic neural network. Samples can be filtered to reduce the ambiguous cases by requiring higher mean probability from the BNN. Using Tensorflow Probability, we developed several BNNs. Our custom model converted to a BNN performed adequately for raw cutouts, but it would not train well for the masked/normalized nuclei. We partially converted Xception to utilize Flipout nodes, leaving the separable convolutions as point estimate nodes. We also fully converted InceptionV3 as an alternative model. Our partial BNN of Xception produced an f1-score of 84%, accuracy of 86%, and AUC of 0.92 (Fig. S2j, k). The full BNN for InceptionV3 gave an f1-score of 79%, accuracy of 80%, and
AUC of 0.87 (Fig. S2l, m). The BNN models can thus be used to understand the probability distribution of the data but at a lower accuracy.

A Deep Neural Network Ensemble Increases Predictive Power

After training the senescent classifier through different sessions, we saw variance in the predictions for a subset of samples. Exploring a large multidimensional solution space during training, neural networks select a relatively good solution that is often biased to favor certain classes 29. Using an ensemble of deep models, the predictions can be combined as though consulting a collection of experts (or interpreted as the “wisdom of the crowd”). To achieve this, we trained an ensemble with random initial weights, potentially allowing convergence to different local minima. We found that there is consistent agreement for the majority of samples, however, there is a significant percent of edge cases with a high variance in predictions among the model instances (Fig. 2a).

We therefore speculate that using an ensemble of deep models for inference and aggregating the results provides predictions with less bias and higher confidence (Fig. 2b). Evidently, some models balance the accuracy of each class in the middle of the range (75-80%), while other models skew toward one class at the expense of the other (for example, obtaining ~85% on one but ~70% on the other). While ensembles have benefits like a BNN, they can be less biased since each ensemble member can specialize around a solution, while a BNN is confined to a single local minima in solution space. Accordingly, we obtained good results with the ensemble method, with an f1-score of 91%, accuracy of 94%, AUC of 0.98 (Fig. 2c, d). More importantly, the ensemble provides a higher confidence and less biased approach by combining multiple models that specialize in predicting different classes.
Figure 2 Predictions from deep ensembles increases accuracy. 

a) Heatmap of variation in predictions by members of ensemble (500 sample nuclei as rows, ensemble members as columns). Blue is young/control and white is senescent.

b) Heatmap of per-class accuracy for control and senescent by ensemble model.

c) Accuracy of deep ensemble.

d) ROC curve for the deep ensemble.

e) Accuracy of single model, Bayesian neural networks, deep ensemble, and bagging.

f) Accuracy of deep ensemble with normalized samples.

g) ROC curve for the deep ensemble with normalized samples.

h) Accuracy of three-state senescence ensemble model.

i) ROC curve for the three-state senescence ensemble model.

j) Accuracy of RS-only model.

k) Accuracy of IR-only model.
An ensemble of neural networks outperforms Bayesian neural networks

We also tried Bagging, where bootstrapping with replacement selects a subset of the samples to use in training independent models. This method did not provide a significant improvement over the basic deep ensemble method (Fig. 2e). The BNN models can be used to improve confidence but sacrifice performance, while the ensemble models provide both (Fig. 2e). We therefore further evaluated the deep ensemble method with masked and normalized samples. This produced an f1-score of 80%, accuracy of 82%, and AUC of 0.89 (Fig. 2f, g), which improved upon the single model. The ensemble method was also applied to the tri-state model to distinguish senescent type, which achieved overall accuracy of 66% and AUC of 0.81 for RS and 0.92 for IR (Fig. 2h, i). While this is lower accuracy, it is an overall improvement of 23.64% compared to the single normalized tri-state model. With all states well above the 33.3% accuracy expected from random predictions, this model is capable of recognizing type of senescence given an adequate sample size.

Due to the lower performance of senescent type prediction, we trained deep models on each type of senescence exclusively, training for control vs RS-only and control vs IR-only. This leaves the other state undefined, assessing each type of senescence separately. Both models classified IR with high accuracy, but the RS-only model recognized RS with ~13% higher accuracy, while the IR-only misclassified those as control (Fig. 2j, k). Ensembles of deep neural networks clearly allow for greater accuracy for senescence prediction.
Modifying Thresholds Increases the Accuracy of Prediction and Improves Confidence

Deep neural networks utilizing one-hot node outputs with the softmax function are trained to produce numerical values that are sometimes treated as the probability for each state. They should not be interpreted as model confidence, but by sampling from a BNN or deep ensemble, we can utilize the distribution to determine uncertainty. We evaluated the predictions for the BNN and deep ensemble (Fig. S3a, b). Correct predictions are indeed oriented toward the lower and higher range of the softmax output, representing greater certainty about a sample’s state. In both cases, the incorrect predictions are clustered toward the center near the 0.5 threshold. Different models could be biased toward either state by shifting those ambiguous samples across the threshold.

We can assume higher confidence in a model’s predictions by raising the classification threshold (of both one-hot states, thereby filtering the predictions in the middle). We therefore evaluated the accuracy using a range of thresholds from 0.5 up to 0.95 in the single model, the Xception BNN, the ensemble of models, and the ensemble of fully normalized models (Fig. S3c, d, e, f). In all cases, we see a significant increase in accuracy as the threshold is raised, due to the ambiguous samples being discarded. By raising the threshold, the Xception-based BNN goes from 85.6% to 96.0%, while the ensemble of normalized models goes from 81.6% accuracy to 97.2%. A similar approach was applied to other models, including the IR-only and RS-only models (Fig. S3g, h). Raising the threshold, these also showed a gain in accuracy of 10-15%. Unfortunately, this led to a significant reduction in the number of samples considered. There is a tradeoff between number of predictions and accuracy, which must be balanced for each application to ensure suitable power for analysis.
Predictor Tracks Development of the Senescent Phenotype

To better understand the development of the senescent phenotype and how nuclear morphology changes over time, we analyzed fibroblasts induced to senescence by 10 Gys of IR and imaged at several time points, including 10, 17, 24, and 31 days. The predictor identifies senescence at all four times points with probability that increases from days 10 to 17 but declines by day 31 (Fig. S4a). Interestingly, examining the probability distribution of the predictor it was apparent that a growing peak of non-senescent cells appear after day 17, either suggesting that the predictor is unable to accurately predict those cells or that a small number of cells may have escaped senescence and are eventually overgrowing the non-senescent cells (Fig. S4b). Indeed, when investigating markers of proliferation, we see that over the time course, PCNA declines until day 17 after which the expression starts to return (Fig. S4c). p21 follows an inverse pattern with stain intensity increasing initially and then declining slightly by day 31 (Fig. S4d). We also saw a decrease in DAPI intensity for days 10 and 17, indicating senescence, but a reversion to control level by day 31 (Fig. S4e). To confirm that the predictor accurately determined senescence even 31 days after IR, we evaluated if markers of proliferation and senescence correlated with predicted senescence. Accordingly, cells with predicted senescence had higher p21 levels, lower PCNA and lower DAPI intensities and vice versa (Fig. S4f, g, h). Morphologically, area and aspect are higher for predicted senescence while convexity is lower (Fig. S4i, j, k). Finally, a simple nuclei count confirms growth, following IR treatment (Fig. S4l). Overall, the senescence predictor captures the state during development in agreement with multiple markers and morphological signs.
DNA Damage Foci and Area Correlates with Senescent Prediction

Senescent cells are associated with the appearance of persistent nuclear foci of the DNA damage markers γH2AX and 53BP1. We characterized the DNA damage foci for our cell lines and investigated how these foci relate to predicted senescence. Our base data set including control, RS, and IR lines were examined for damage foci. Using high content microscopy, we counted DNA damage foci per nuclei and found the mean count of γH2AX and 53BP1 foci to be below 1 each (0.9 and 0.6, respectively) for controls, while RS had 4.0 γH2AX and 2.0 53BP1 foci and IR had 3.4 γH2AX and 3.0 53BP1 foci (Fig. 3a, b, S5a). To study how the presence of damage foci relates to predicted senescence, we calculated the Pearson Correlation between predicted senescence and γH2AX and 53BP1 foci counts. We found that across all conditions there is a moderately strong correlation of around 0.5 (Fig. 3c). This association is also visible when simply plotting foci counts and senescence prediction which shows predicted senescence flipping from low to high, along with shifts in foci counts (Fig. S5b). Within senescent subtypes RS and IR, the correlation is slightly weaker, perhaps indicating that the senescent probability score for each subtype has some correlation with foci count. Our feature reduction including masking means that internal nuclear structure was not used in assessment, but it is nonetheless notable that senescence prediction (overall and by subtype) correlates with foci count. We also compared the correlation between predicted senescence and area. Here too, we see a correlation of around 0.5, and slightly weaker for the subtypes. In sum, there is a considerable correlation between foci counts and senescence.
Figure 3 Senescence can be predicted across tissues and species. a Number of γH2AX foci by type of senescence (n=1,831-15,560, mean ± 95% CI). b Number of 53BP1 foci by type of senescence (n=1,831-15,560, mean ± 95% CI). c Correlation between foci count and predicted senescence. d Representative immunohistochemistry micrographs of premature aging nuclei with DNA damage foci staining of γH2AX and 53BP1, HGPS: Hutchinson-Gilford Progeria Syndrome, AT: ataxia telangiectasia, CS: Cockayne Syndrome. e Nuclear area for premature aging diseases (n=4,340-15,074, mean ± 95% CI), HGPS: Hutchinson-Gilford Progeria Syndrome, AT: ataxia telangiectasia, CS: Cockayne Syndrome. f Number of γH2AX foci for premature aging diseases (n=5,162-17,584, mean ± 95% CI). g Number of 53BP1 foci by premature aging diseases (n=5,162-17,584, mean ± 95% CI). h Predicted probability of senescence for premature aging disease (n=5,162-17,584, mean ± 95% CI). i Nuclear area of murine astrocytes (n=4,918-13,661, mean ± 95% CI). j Representative immunohistochemistry micrographs of senescent murine astrocytes with DNA damage foci staining of γH2AX and 53BP1. k DAPI intensities for premature aging diseases and controls (n=4,340-15,074, mean ± 95% CI). l Nuclear area of murine astrocytes (n=4,888-13,549, mean ± 95% CI). m Nuclear area of murine neurons (n=33,303-62,847, mean ± 95% CI). n Nuclear area of murine neurons (n=33,303-62,847, mean ± 95% CI). o Number of γH2AX foci for murine astrocytes (n=4,918-13,661, mean ± 95% CI). p Number of 53BP1 foci for murine astrocytes (n=4,918-13,661, mean ± 95% CI). q Number of γH2AX foci for murine astrocytes (n=33,303-62,847, mean ± 95% CI). r Number of 53BP1 foci for murine astrocytes (n=33,303-62,847, mean ± 95% CI). s Predicted senescence for murine astrocytes (n=4,918-13,661, mean ± 95% CI). t Predicted senescence for murine neurons (n=33,303-62,847, mean ± 95% CI).
Patients suffering from premature aging, or progeria, represent genetically well-defined models to understand the molecular basis of aging\textsuperscript{32,33}. To test if cell lines from progeria patients display accelerated aging in culture, we applied the senescent classifier to primary fibroblasts isolated from Hutchinson-Gilford progeria syndrome (HGPS), ataxia telangiectasia (AT) and Cockayne syndrome (CS) (Fig. 3d). Evaluating the area of the nuclei of progeria cells, we found that in general their mean is significantly larger than controls. Notably ataxia-telangiectasia cells have the largest nuclei at 25% higher than controls, while Hutchinson-Gilford progeria and Cockayne syndrome are both 15% higher (Fig. 3e). We also investigated DNA damage foci and observe that most prematurely aged lines have higher γH2AX and 53BP1 foci counts (Fig. 3f, g, S5c). Further, despite diverse mechanisms, the classifier recognized these cell lines having significantly greater probability of senescence (Fig. 3h). All progeria lines have high mean probability of senescence at 0.7, indicating that the average cell in each group is considered senescent, while controls are below the standard threshold at 0.3. Evaluating SA-β-gal activity, we find 35-60% of nuclei have positivity and overall correlation of 0.5 between predicted senescence and having nearby SA-β-gal (Fig. 3i). When predictions are filtered to higher confidence levels, there is an increase in correlation up to 0.9 (Fig. 3j), indicating high confidence predictions are capturing the senescent state. DAPI intensity also suggests that all progeria lines have higher senescence compared to controls (Fig. 3k). These observations indicate that our classifier may be able to discriminate rates of aging in cultured cells.
The senescent classifier translates across species and cell types in culture

To broaden the applicability of our classifier we speculated that it might apply to nuclei from other cell lines and species. We therefore evaluated the model on mouse primary astrocytes and neurons treated with IR (Fig. 3l). While astrocytes are known to senesce with cell cycle arrest, post-mitotic neurons also exhibit a senescence-like state. We first compared the nuclei area and found that the IR-treated astrocytes had slightly but significantly larger nuclei than controls while IR-treated neurons had reduced area, unlike other cell types we studied (Fig. 3m, n). Evaluating DNA damage foci, we see that IR treated astrocytes and neurons have substantially higher foci count as expected (Fig. 3o, p, q, r). We next applied the ensemble of deep models and found that the IR treated astrocytes had a 7.7% higher probability of senescence than controls while IR-treated neurons have 6.3% higher probability (Fig. 3s, t).

Senescence prediction translates across species and tissues in vivo

We applied the predictor to H&E-stained liver tissue from C57Bl6 mice at taken at 48, 58, and 78 weeks of age. After imaging the tissue sections at 20x, we used a deep learning segmentation model trained on 18 tiles to extract nuclei from 16,187 tiles (Fig. 4a). Our training set included samples of hepatocytes only, and this cell type was primarily selected during automated segmentation. We first analyzed morphological metrics, finding an insignificant increase in nuclear area (Fig. 4b). However, we saw a significant decrease in convexity and increase in aspect ratio, both indicating increased senescence with age (Fig. 4c, d). Nuclei were evaluated for senescence using the normalized RS-only and IR-only models, of which the RS model indicated increasing senescence with age while the IR model did not (Fig. 4e, f). Using the probability, we calculated the percent of senescent cells, finding ~36% for RS and...
Figure 4 Senescence can be predicted across tissues and species. a Analysis workflow. b Mean nuclear area per mouse by age (n=5). c Mean nuclear convexity per mouse by age (n=5). d Mean nuclear aspect ratio per mouse by age (n=5). e Prediction probability for RS senescent (n=5). f Prediction probability for IR senescent (n=5). g Predicted percent that are RS senescent (n=5). h Predicted percent that are IR senescent (n=5). i Mean probability of predicted RS senescence for each individual by p21 state. j Mean probability of predicted IR senescence for each individual by p21 state. k PCNA intensity after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). l DAPI intensity after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). m Aspect after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). n DAPI intensity after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). o Predicted probability of senescence after senescence induction, using the RS model (n=30,957-119,669, mean ± 95% CI). p Accuracy of doxorubicin-only model. q Accuracy of ATV/r-only model. r Accuracy of Antimycin-A-only model. s Accuracy of unified model.
The predictor is trained on DAPI-stained cultured fibroblasts representing a considerable difference in context, it is therefore likely that the algorithm should be tuned to evaluate other data sources. Applying thresholds of 0.6 and 0.9 for RS and IR, respectively, the percent was brought down to roughly 8-10% to match the percent reported, roughly adjusted for difference in age and split between IR and RS. With these thresholds, the percent of senescent cells per mouse increased with age. To determine the predictor’s ability to identify senescent hepatocytes in liver tissue, we also stained tissue sections from the same specimens with DAPI and p21, identified hepatocytes with segmentation, and predicted senescence of those nuclei. We found that the mean predicted senescence per animal for the p21+ cells was significantly higher compared to p21- for both RS and IR models. Given the differences in human and mouse nuclei as well as between cell types, it is notable that the senescent state can be captured through the relative difference in assessed probability. It therefore appears that our predictor may be able to determine senescence across cell types and species.

Predictor to Recognize Senescence from Multiple Mechanisms

Senescence can be induced by several types of stressors that could result in different types of pathologies. We therefore induced senescence by different drug treatments, including the DNA damaging compound doxorubicin, the mitochondrial toxin antimycin A, and the HIV protease inhibitor atazanavir/ritonavir (ATV/r) and evaluated predictors for each of them. PCNA staining confirmed that each drug treatment led to a senescent state. A morphometric analysis showed that all three drug treatments expanded nuclear area, decreased convexity, and increased aspect ratio. Additionally, DAPI intensity decreased significantly for all three treatments, indicating senescence. The predictor model (trained on IR and RS...
recognized senescence in the nuclei treated with doxorubicin but did not detect senescence in treatment with antimycin A or ATV/r (Fig. 4p). We speculate that doxorubicin treatment more closely resembles the DNA damage caused by IR-induced and replicative senescence. To address this limitation of our model, we trained new models for each new type, including doxorubicin-only, Antimycin A-only, and ATV/r-only (Fig. 4q, r, s). In addition, we trained on a broader data set, including IR, RS, doxorubicin, antimycin A, and ATV/r. Tested on validation data held out from training, we find the expanded model can now recognize antimycin A with 66.0% accuracy, ATV/r with 64.3% accuracy, and doxorubicin with 62.3% accuracy, which exceeds performance for each individual predictor (Fig. 4t). However, it has reduced accuracy for IR at 68.1% (compared to 83.0%), although RS is slightly higher in this model. While the base predictor model provides higher accuracy for IR, the unified model can recognize senescence in more diverse conditions.

**Predictor Detects Senescence in Human Dermal Tissue with p21**

To determine if the predictor could be used with human dermis, we analyzed samples from an independent data set, stained with hematoxylin and DAB for p21 (Fig. S6a). Nuclei were detected using image segmentation with U-NET trained on the hematoxylin nuclei, and the predictor generated senescence probability scores for the extracted nuclei. After calibrating the p21 detection threshold to roughly match published rates, we found the mean predicted senescence of p21+ nuclei was 5.9% higher than those without p21 for the RS and IR models (with p=0.005 for both), while other models showed no difference (Fig. 5a, S6b, c, d, e). As the confidence threshold was raised above the standard 0.5, p21+ was clearly separated from p21-. With increasing confidence, the p21+ nuclei generally showed higher predicted probability for IR, while the p21- nuclei showed lower predicted probability for RS. The percent difference
between mean p21+ and p21- probability also increased with higher confidence. Notably, all
three other models (for doxorubicin, ATV/r, and antimycin-A) showed no separation between
p21+ and p21-, indicating that they are picking up on other type-specific aspects of senescence.

The human dermis shows age-dependent increase in senescent nuclei

To further investigate if the predictor could be applied in a clinical context, we tested the
algorithm on human skin samples of 169 individuals aged 20-86 years. The senescent classifier
was used to evaluate the dermal nuclei from biopsy samples, stained with H&E and imaged in a
slide scanner at 20x. We used U-Net to detect nuclei, extracted nuclear regions, and converted
the nuclei to the normalized and masked form (Fig. 5b). We first evaluated several
morphological metrics, including area, convexity, and aspect ratio. Across age, we see no
change in area (Fig. 5c), an insignificant change in convexity (Fig. 5d), and a significant change
in aspect ratio (Fig. 5e). We considered that different pathologies could be related to various
forms of senescence (senescence caused by diverse mechanisms such as DNA damage,
telomere attrition, mitochondrial dysfunction, and so on), so we evaluated multiple senescence
predictor models developed here. We found the probability of senescence increases with age of
patients for RS but is relatively flat for IR and declines for ATV/r, antimycin-A and doxorubicin
(Figures S6f, g, h, i, j). We expect a percent of human dermal nuclei to be senescent, ranging
from ~1% in young to ~15% in old, so we selected thresholds to calibrate the model with 0.7
for RS and 0.85 for IR, leading to an overall predicted percent of ~6% and showing an age-
dependent increase in percent of senescence (Fig. 5f, g). Both IR and RS models predict a
statistically significant increase with age, while doxorubicin, Antimycin-A, and ATV/r appear
decoupled from age (Fig. S6k, l, m). We also evaluated the correlation between morphological
metrics and predicted senescence and found moderate correlation for several metrics, but RS
Figure 5: Nuclear morphology predict senescence and multiple diseases in humans. 

- **a** Predicted probability of senescence for p21- and p21+ nuclei in human dermis, across a range of thresholds.
- **b** Analysis workflow.
- **c** Mean nuclear area per patient by age (n=148).
- **d** Mean nuclear convexity per patient by age (n=148).
- **e** Predicted percent that are RS senescent (n=169).
- **f** Predicted percent that are IR senescent (n=169).
- **g** Volcano plot of ICD-10 chapters based on IR senescence residuals and chi-squared p-values.
- **h** Volcano plot of ICD-10 chapters based on RS senescence residuals and chi-squared p-values.
- **i** Volcano plot of ICD-10 chapters based on doxorubicin senescence residuals and chi-squared p-values.
- **j** Volcano plot of ICD-10 chapters based on ATV/r senescence residuals and chi-squared p-values.
- **k** Volcano plot of ICD-10 chapters based on antimycin-A senescence residuals and chi-squared p-values.
- **l** Volcano plot of ICD-10 chapters based on unified senescence residuals and chi-squared p-values.
- **m** Table summarizing disease conditions, percent of individuals with the condition for positive and negative residual groups, p-value from chi-squared test or Fisher’s exact test, and relative risk ratio with 95% CI when significant.
was more correlated with convexity while IR was more correlated with area and aspect ratio, perhaps indicating morphological aspects of each type of senescence in vivo (Fig. S6n, S7a).

Interestingly, we found that area was anti-correlated with both predicted IR and RS, but predicted IR was inverse to aspect ratio. This indicates difference between senescence in culture and in tissue sections and affirms that the IR and RS model are picking up on different aspects of senescence. We considered whether the age-dependent increase in predicted senescence could be related to change in proportions of detected cell types, so we compared the distribution of cell area and aspect ratio for broad age groups, individuals below 40 and those over 60. A shift in cell types should be reflected by a change in these metrics, but we found no noticeable difference in the distribution of these metrics for predicted senescent and non-senescent cells between age groups (Fig. S7b, c). Comparing each group by mean area and aspect ratio of individuals, a t-test shows non-significance (p=0.94 and p=0.51, respectively), indicating that each group has a similar proportion of cell types.

Senescent dermal nuclei are inversely associated with neoplasms and positively correlated with hypertension and osteoporosis

Given the large variation in predicted senescence, we speculated that these values could represent meaningful health outcomes. To investigate, we retrieved 19,820 ICD-10 diagnosis codes collected in the Danish National Patient Register from 1977 to 2018 for all the individuals in the study. We looked for associations between individuals with diagnosed conditions grouped by ICD-10 chapters and predicted senescence above or below the age-dependent mean (those above or below the trendline in Fig. 5f, g and S6f-m specifically using residuals from linear regression of predicted senescence versus age), using the chi-squared test and Fisher’s exact test for the frequency of occurrence between the two groups (Fig. 5h-m). Remarkably, we found
a significant correlation between a rate of senescence below the age-matched mean and the presence of ICD-10 Chapter II Neoplasm diagnosis codes for both RS and IR, with p-values of 0.002 and 0.005, respectively (Fig. 5n). Narrowing down the analysis we determined the association was based on malignant (versus benign or unknown) codes within ICD-10 Chapter II Neoplasm with IR p-value at 0.018 and RS at 0.058. Notably, RS better represents replicative senescence which occurs naturally with age, while IR better represents DNA damage, although there is considerable overlap in predictions between the two with this model. We also scanned individual ICD-10 clinical codes and found several other conditions associated with senescence, including osteoporosis, osteoarthritis, hypertension, cerebral infarction, hyperlipidemia, hypercholesteremia, and hearing loss, which were all significant when evaluated individually but non-significant when applying multiple test correction, such as the Bonferroni (Fig. 5n, S7d-o, S8a). All of these conditions were associated with higher levels of predicted dermal senescence except for cancer and hearing loss, which were associated with lower levels of predicted senescence. They draw from different models, for example neoplasms are particularly significant with RS, while hypertension only appear in the Antimycin-A model. Overall, we found that high assessed senescence corresponds to fewer neoplasms and malignancies, while also indicating increased frequency of osteoporosis, osteoarthritis, hypertension, and other conditions.
In this paper we present a neural network classifier that can predict cellular senescence based on nuclear morphology. Trained on fibroblasts maintained in cell culture, the classifier achieves very accurate results, which was confirmed by applying it to independent cell lines. We also trained models to correctly distinguish between senescence caused by radiation induced damage and replicative exhaustion. By training additional models on samples with reduced features, we infer that the shape of the nucleus alone provides a significant signal to indicate senescent state. DAPI-stained nuclei with background removed, size normalized, and internal structure masked are still classified with high accuracy. These feature reduction methods serve a secondary purpose, making a model robust to technical variation – our neural network trained on reduced samples can make predictions on nuclei that were prepared in other experimental and imaging contexts. Indeed, the predictor distinguished senescent astrocytes and neurons, predicted an age-related increase in senescent liver cells, and confirmed senescence in cell lines from patients suffering from premature aging. Although it is still debated if universal markers of senescence exist, our findings suggest that at least morphological alterations in nuclei may be common across some tissues and species.

We present several predictor models, including those that combine IR, RS and other methods, and those that specialize on each for improved accuracy. The base model trained on IR and RS can identify either type along with senescence induced by doxorubicin, indicating that the predictor has identified features found in multiple types related to DNA damage. Our base model did not accurately identify ATV/r and antimycin A, but a new model trained on all five methods could accurately identify senescence induced by these diverse mechanisms. The
unified model could be identifying a common signature or simply recognizing multiple phenotypes.

Our data shows that individuals with a predicted higher rate of senescent cells have reduced neoplasms and malignant cancer, in comparison to those with a lower rate of senescence. This is highly consistent with the notion that senescence is a likely mechanism to control cancer development by limiting uncontrolled proliferation. Further, premalignant tumors express markers of senescence, which are absent in malignancies, and malignant tumors can regress and undergo senescence by switching off oncogenes, supporting the protective role of senescence in blocking the progression of neoplasms to malignancies. In addition, loss of central senescence inducers such as p16 is very common in many cancer types. Of note, there is also evidence suggesting that cellular senescence promotes malignancy through the inflammatory senescence associated secretory phenotype or SASP, that senescent cells may appear in areas where tumors tend to subsequently develop, and that senescent cells and SASP induced by cancer treatment led to worse survival and healthspan. While the role of senescence in cancer is highly complex, our results based on clinical data support the overall protective role for senescence in human health with regards to cancer. We also found several other conditions often associated with senescence, including osteoporosis, osteoarthritis, hypertension, cerebral infarction, hyperlipidemia, and hypercholesteremia, which appear more frequently in individuals with a higher predicted rate of senescence.

We also investigated how our deep learning predictor results correspond to other measures of senescence. Nuclear area is known to expand during senescence, and we confirmed this fact in our cell culture data set, with significant differences in IR and RS senescent cells. On a per nuclei basis, we found a moderate correlation between area and predicted senescence. However, due to our size normalization, it is unlikely this classic feature is the primary signal for
our deep learning model (at least for the size-normalized version). We also identified convexity and aspect ratio as key morphological properties that differ between control and senescent cells in culture and found moderate correlation between each of these properties and predicted senescence. Interestingly, we found no increase in area with age in the human dermis, but a significant increase in aspect ratio and significant decrease in convexity, indicating nuclei becoming stretched and irregular with advancing age in humans. These observations confirm that size normalization is necessary to generalize our neural network classifier. It also demonstrates the value of our feature-neutral approach, where the neural network is trained to identify senescence from rich image data, and it is later reduced through feature removal.

In sum, our deep neural network model is capable of accurately predicting the senescent state and type from nuclear morphology using several imaging techniques and has been demonstrated with several diverse applications. We applied the predictor to human skin samples and observed an age-dependent increase in senescence. Remarkably, individuals who appear to have higher rates of senescent cells show reduced incidence of malignant neoplasms. This supports the long-standing hypothesis that senescence is a mechanism to limit cancer. Further, we find association between higher predicted senescent cell burden and other conditions, including osteoporosis, osteoarthritis, hypertension, cerebral infarction, hyperlipidemia, and hypercholesteremia.
Methods

Cell culture

All human-derived primary skin fibroblast cells were purchased from Coriell Institute (USA).

Control fibroblasts included AG08498 (male, 1 year), GM22159 (male, 1 day), GM22222 (male 1 day), GM03349 (male, 10 years) and GM05757 (male, 7 years). Cells were cultured at 37°C and 5% CO2 either in 1:1 mix of DMEM GlutaMAX (Gibco, 31966047) and F-12 media (Gibco, 31765068) for AG08498, GM22159 and GM22222 or in EMEM media (Biowest, L0415-500) for GM03349 and GM05757. Fibroblasts derived from Hutchinson-Gilford progeria syndrome patients included AG06917 (male, 3 years), AG06297 (male, 8 years) and AG11513 (female, 8 years). Fibroblasts sampled from ataxia telangiectasia and Cockayne syndrome patients were GM03395 (male, 13 years) and GM01428 (female, 8 years), correspondingly. Cells were cultured at 37°C and 5% CO2 in MEM media (Lonza, BE12-662F). Freshly isolated primary mouse astrocytes were kindly provided by the Department of Drug Design and Pharmacology, University of Copenhagen. Cells were cultured at 37°C and 5% CO2 in DMEM GlutaMAX (Gibco, 31966047). All used media were supplemented with 10% fetal bovine serum (Sigma-Aldrich, F9665) and 100 U/mL penicillin-streptomycin (Gibco, 15140163).

Senescence induction

To achieve replicative senescence control fibroblasts at early passages were seeded in T25 cell culture flasks (200 000 cells) and cultured over 32 weeks. After each splitting cell number was recorded and population doubling level (PDL) was calculated as \( \log_2(\text{cell number during harvesting/\text{cell number during seeding}}) \). Experiment was terminated when PDL reached zero.

Induction of cellular senescence by ionizing radiation, doxorubicin, antimycin A and atazanavir/ritonavir was performed according to 46. Briefly, control fibroblast cells at early
passages were seeded in 96 well plates (Corning, 3340) in a density of 2 000 cells per well. Day after cells either were exposed to 10Gy of ionizing radiation or treated with 250 nM doxorubicin for 24h and cultured for the next nine days. Medium was replaced every two days. Three days before radiated or doxorubicin-treated cells reached senescence state, fibroblast cells from the same stock were seeded (2 000 cells/well) as mock-irradiated or DMSO-treated controls. Mitochondrial dysfunction-induced senescence was achieved by treating control fibroblast cells with 250 nM antimycin A every two days within ten days. 25 μM atazanavir/ritonavir was given to control fibroblast cells every two days within fourteen days to develop senescence phenotype. Corresponding DMSO-treated controls were cultured in parallel and seeded in 96-well plate three days before terminating the experiment..

**Gene expression analysis**

Radiated and control fibroblast cells were lysed using TRizol reagent (Ambion by Life technologies) and RNA was isolated using Direct-zol RNA miniprep Plus (Zymo Research). RNA was reverse transcribed into cDNA using High-capacity cDNA reverse transcription kit (Applied Biosciences) according to manufacturer’s protocols. Next, RT-qPCR was performed to detect mRNA levels of senescence markers using StepOnePlus Real-Time PCR System (Applied Biosystems). Relative gene expression was calculated based on obtained Ct values, normalized to housekeeping gene GAPDH, and expressed as fold changes compared to non-irradiated control. The following specific primer sequences were used:

- **CDKN2A (p16):**
  - Forward: GAG CAG CAT GGA GCC TTC
  - Reverse: CGT AAC TAT TCG GTG CGT TG

- **CDKN1A (p21):**
  - Forward: TCA CTG TCT TGT ACC CTT GTG C
Reverse: GGC GTT TGG AGT GGT AGA AA

IL6:
Forward: CAG GAG CCC AGC TAT GAA CT
Reverse: GAA GGC AGC AGG CAA CAC
GAPDH:
Forward: GTC AGC CGC ATC TTC TTT TG
Reverse: GCG CCC AAT ACG ACC AAA TC

In total three independent sets were performed in GM22159 cells.

**Immunocytochemistry, SA-β-galactosidase detection and image preparation**

For detection of persistent DNA damage foci as well as fluorescence intensity levels of p16, p21, p53 and PCNA, fibroblast cells were washed once with warm PBS, fixed in 4% paraformaldehyde (PFA) for 15 min followed by permeabilization step with incubation for 10 min in PBS-0.1% Triton X100. Blocking was performed in 1% BSA-PBS-0.1% Tween 20 overnight at 4C. Next day cells were incubated with primary antibodies (γH2AX, 1:1000, Millipore, 05-636 and 53BP1, 1:2000, Novus, NB100-304; p16\[^{INK4A}\], 1:50, Santa Cruz, sc-56330; p21, 1:200, Santa Cruz, sc-6246; p53, 1:200, Santa Cruz, sc-126; PCNA, 1:500, Abcam, ab18197) for 1h at RT, washed three times with PBST and incubated with secondary antibodies (1:200 Alexa-Flour 488, Invitrogen, 10424752 and 1:200 Alexa-Flour-568, Invitrogen, 10348072) for 1h at RT. Cells were incubated with DAPI solution (AppliChem, A4099) for 10 min and stored in PBS at 4C until the analysis.

SA-β-gal was detected using senescence cells histochemical staining kit (Sigma-Aldrich, CS0030) according to manufacturer's protocol. Cell colonies were imaged using INcell analyzer 2200 high content microscopy at 20x magnification to produce 1199 images with 2048x2048
pixel resolution. Due to system constraints for object detection, each image was split into four
tiles of 1024x1024 pixel resolution.

Cell Cycle Analysis

Fibroblast cell lines AG08498, GM22159 and GM22222 were seeded in 6-well cell culture plates
and 24 hours later DNA-damage induced senescence was conducted as described above (5
Gy, 10 Gy or 20 Gy). After 9 days cells were harvested by trypsinization and washed twice with
PBS. Cells were fixed by adding dropwise ice-cold 70% ethanol while mixing the cells gently on
a vortex mixer. Thereafter, the cells were incubated for 30 min on ice and washed twice in PBS.
Fixed cells were incubated with RNase (100 µg/ml; ThermoFisher Scientific) at 37°C for 30 min.
Propidium iodide (20 µg/ml; Sigma-Aldrich) was added and incubated for 30 min at 37°C. Cell
cycle status was determined by flow cytometry (CytoFlex, Beckman Coulter).

Nuclei Detection

A base library was prepared using control, irradiated (IR), and cells serially passaged until they
reached senescence (replicative senescence, RS). A deep neural network model was applied to
detect DAPI-stained nuclei. The samples were used to build a training set for nuclei recognition.
Several images were selected arbitrarily from each group for a total of ~20 samples, and all
nuclei in the training samples were annotated by selecting the nuclear region. U-NET, a 23-layer
fully convolutional network for image segmentation, was trained using the samples, learning to
associate the DAPI images with annotation masks indicating nuclear regions. Our
implementation of U-NET is largely based on the original U-NET\footnote{47}, but includes a dropout layer
after each of the convolutional and deconvolutional layers to reduce overfitting. After training for
1000 epochs, the U-NET model was used to detect nuclei for all 4796 tiles (1199 images x 4
tiles/image), producing output images of predicted nuclei regions. The images with predicted
nuclei were scanned for recognition regions of area between 500 and 15,000 pixels. Each
detected nucleus was extracted along with its surrounding context as a centered 128x128 pixel region and used to assemble a base library of 95,152 nuclei. In addition, the recognition region itself was cutout, providing a two-color reduction of the detected nuclei, and assembled into a secondary library of nuclei masks.

Nuclear Morphology

An analysis of the nuclei was performed to assess morphological properties. The two-color mask library was used, since it provided a universal representation of the detected nuclei (with U-NET detector models that have good coverage of the nuclei region). Nuclear morphology was assessed using several metrics, including area, perimeter, moments, convexity, and aspect ratio. Convexity is the ratio of perimeter to convex hull perimeter, which provides a size-neutral measure of boundary regularity. The convex hull is a polygon that connects the outer edges of nuclei like an envelope.

Senescent Classification

After assembling a library of senescent cells, a deep neural network was trained to classify DAPI-stained nuclei as senescent or non-senescent. The training set was based on several cell lines GM22159, GM03349, GM05757, while additional cell lines GM22222 and AG08498 were used for testing. Training samples were randomized and split into 80% for training and 20% for validation. Due to experimental setup, the sample classes are unbalanced, with 75.2% control, 11.2% RS, and 13.6% IR. The samples were balanced during training by applying class weights with inverse proportion to the class abundance (for example, senescent samples composed of IR and RS were fewer in number and therefore valued 3x higher than controls). Image samples were normalized for brightness/intensity by adjusting each image’s mean intensity to 0 and standard deviation to 1. Augmentation was also applied during training, randomly modifying samples: adjusting size from 80% to 120%, changing normalized brightness from 70% to 130%,
flipping horizontally and vertically, and rotating up to 180 degrees. For each epoch, one augmentation cycle was performed. Training was done with Xception, a 48-layer model, initialized with ImageNet weights but set to allow weight adjustment of all layers during training. The top layer was replaced by a layer of one-hot nodes to indicate the state as controls or senescent (or as a tri-state model with controls, IR, or RS to indicate the type of senescence).

With this minor adjustment, the model provided 37,640,234 trainable parameters. Training was done using Adam with the learning rate set to $1 \times 10^{-4}$ for 10 epochs, in which time accuracy rapidly converged to a steady level. In addition, a simpler custom model was tested, with three convolutional layers with ReLU activation and two dense layers with L1/L2 regularization of 0.05/0.05 and 30% dropout. This model required 713,296 parameters. For both network designs, we trained with raw images along with several modified image sets, where the background was removed, the nuclei were size normalized, and the inner details of nuclei were entirely masked (Fig. 1A). All three techniques were based on the detected nuclei. To remove the background, the area outside of the nuclei was set to 0. Size was normalized by rescaling all nuclei so the larger of the two dimensions was a standard size of 80 pixels. Finally, the size-normalized detection region was used for the masked nuclei set.

**Bayesian Neural Network**

We used Tensorflow Probability to create a Bayesian neural network (BNN). We first converted the simple custom model, replacing nodes with the comparable FlipOut version, which assumes that the kernel and bias are drawn from a normal distribution. During a forward pass, kernels and biases are sampled from posterior distribution. Targets were encoded as above, and the loss function used was cross entropy plus KL divergence divided by number of batches. We also partially converted Xception to a BNN by replacing all dense and convolutional layers to FlipOut nodes, leaving separable convolutions unconverted since a FlipOut version was not available. In addition, we fully converted InceptionV3 for evaluation. Inference was done by
evaluating the model 20 times to produce a distribution of predictions, and then taking the mean probability for each sample.

**Deep Neural Network Ensemble**

To improve accuracy and provide a more robust solution, we also worked with an ensemble of deep learning models. This method utilized 10 models of Xception, each trained on the same data set with different random weight initialization. To generate predictions, each model instance was applied, and the results combined by taking the mean prediction. We also tried bagging, also known as bootstrap aggregation. Similar to the deep ensemble, this method trains different model instances with bootstrap selection of samples for n=1-1/e. With each instance trained on a different subset of samples, this method produces multiple models that in theory can specialize to different sets of data.

**Statistical Methods**

All comparisons with between groups of samples were made using one-way ANOVA f-tests to evaluate differences in the means, followed by pair-wise tests using Tukey’s HSD (Honest Significant Difference) to calculate p-values between groups. Linear regression methods were evaluated with R and p-value statistics. Groups of patients were compared using the chi-squared test and Fisher test to detect significant differences between frequencies. Correlation was evaluated using the Pearson colocalization coefficient.

**Pathology sample selection**

The individuals were sampled from patients for whom samples of naevi on non-sun exposed skin had undergone pathology without malignant findings at a major pathology department in Copenhagen. The patient sample was selected to have flat distribution of age. We selected patient samples from the Danish National Register of Pathology requisitioned in 2007-2017 and
coded with one or more PatoSNOMED topology code: T02530 (Skin on penis), T76330 (Foreskin), T80200 (Mons pubis), T02471 (Skin on nates), T02480 (Skin on abdomen), T02430 (Skin on breasts) and one or more procedure code: P30620 (resect), P306X0 (ectomy preparation), P30611 (excision biopsy) and one or more morphology code: M87400 (junction naevus), M87500 (dermal naevus), M87600 (compound naevus).

Senescence and Human Morbidity

We collected ICD-10 diagnosis codes from the Danish National Patient Register in the period 1977-2018 of each of the patients in this study. We further grouped diagnoses into each of 21 ICD-10 chapters. We calculated the linear regression residuals of the relationship between age at pathology examination and the predicted senescent cell load (IR, RS metrics) for each of the patients. We then constructed contingency tables counting the number of patients with and without a specific diagnosis and with a predicted senescent cell load above or below the age-dependent average. We used Pearson's chi-squared test and Fisher's exact test to determine whether patients with a predicted senescent cell load above or below the age-dependent average were associated with a higher or lower incidence of specific diagnosis codes (or diagnosis within a specific ICD-10 chapter.)

Animals

Male C57BL/6J mice were acquired from Janvier Labs (Le Genest Saint Isle, France). Animals arrived at 5-8 weeks of age and were housed in a controlled environment (12 h light/dark cycle, 21 ± 2 °C, humidity 50 ± 10%). Stratification and randomization into individual diet groups were based on baseline body weight. Mice had ad libitum access to tap water and chow (2018 Teklad Rodent Diet, Envigo, Madison, WI, United States; Altromin 1324, Brogaarden, Hoersholm, Denmark). The study was approved by The Institutional Animal Care and Use Committee at MedImmune (Gaithersburg, MD, United States) and The Danish Animal Experiments
Inspectorate (license: 2017-15-0201-01378) and performed in accordance with internationally accepted principles for the use of laboratory animals.

Liver histology

Terminal liver samples were dissected from the left lateral lobe immediately after sacrificing the animal and subsequently fixed overnight in 4% paraformaldehyde. The liver tissue was then paraffin-embedded and sectioned at a thickness of 3 µm. Sections were stained with hematoxylin-eosin (HE, Dako, Glostrup, Denmark). Slides were scanned by ScanScope AT System (Aperio, Vista, CA, United States).
Acknowledgements

This research was supported by the Novo Nordisk Foundation Challenge Programme (#NNF17OC0027812), the Nordea Foundation (#02-2017-1749), the Neye Foundation, the Lundbeck Foundation (#R324-2019-1492), the Ministry of Higher Education and Science (#0238-00003B), VitaDAO and Insilico Medicine. M.L.I. and M.G. were supported by the NIA IRP, NIH.

Author Contributions

I.H. wrote the article, developed and trained deep learning models, and analyzed data. G.V.M. performed experiments on the base data set, astrocytes, neurons, and premature aging disease. M.B.E. analyzed clinical data. D.B. performed flow experiments. J.S.M. developed Bayesian networks and advised the project. M.H.N. and D.O. performed animal experiments. M.L.I. and M.G. shared images of human dermal tissue from their study and edited the project. L.M. managed clinical images and medical records. E.V. advised and edited the project. R.W. advised and edited the project. M.S.K. conceived the idea, supervised the project and edited the manuscript.

Declaration of Interests

The authors declare no competing interests.
References

1. Kirkland, J. L. & Tchkonia, T. Cellular Senescence: A Translational Perspective. *EBioMedicine* **21**, 21–28 (2017).

2. Zglinicki, T. von, Saretzki, G., Ladhoff, J., Fagagna, F. d’Adda di & Jackson, S. P. Human cell senescence as a DNA damage response. *Mech. Ageing Dev.* **126**, 111–117 (2005).

3. Covarrubias, A. J. *et al.* Senescent cells promote tissue NAD+ decline during ageing via the activation of CD38+ macrophages. *Nat. Metab.* **2**, 1265–1283 (2020).

4. Childs, B. G. *et al.* Senescent cells: an emerging target for diseases of ageing. *Nat. Rev. Drug Discov.* **16**, 718–735 (2017).

5. Schafer, M. J. *et al.* The senescence-associated secretome as an indicator of age and medical risk. *JCI Insight* **5**, e133668 (2020).

6. Young, A. R. J., Narita, M. & Narita, M. Cell Senescence as Both a Dynamic and a Static Phenotype. in *Cell Senescence* (eds. Galluzzi, L., Vitale, I., Kepp, O. & Kroemer, G.) vol. 965 1–13 (Humana Press, 2013).

7. Basisty, N. *et al.* A proteomic atlas of senescence-associated secretomes for aging biomarker development. *PLOS Biol.* **18**, e3000599 (2020).

8. Matjusaitis, M., Chin, G., Sarnoski, E. A. & Stolzing, A. Biomarkers to identify and isolate senescent cells. *Ageing Res. Rev.* **29**, 1–12 (2016).

9. Ogrodnik, M. Cellular aging beyond cellular senescence: Markers of senescence prior to cell cycle arrest *in vitro* and *in vivo*. *Aging Cell* (2021) doi:10.1111/acel.13338.

10. Lee, S. & Schmitt, C. A. The dynamic nature of senescence in cancer. *Nat. Cell Biol.* **21**, 94–101 (2019).

11. Campisi, J. Cellular senescence: putting the paradoxes in perspective. *Curr. Opin. Genet. Dev.* **21**, 107–112 (2011).
12. Gorgoulis, V. et al. Cellular Senescence: Defining a Path Forward. *Cell* **179**, 813–827 (2019).

13. Mitsui, Y. & Schneider, E. L. Increased nuclear sizes in senescent human diploid fibroblast cultures. *Exp. Cell Res.* **100**, 147–152 (1976).

14. Chen, J.-H. & Ozanne, S. E. Deep senescent human fibroblasts show diminished DNA damage foci but retain checkpoint capacity to oxidative stress. *FEBS Lett.* **580**, 6669–6673 (2006).

15. Kusumoto, D. et al. Anti-senescent drug screening by deep learning-based morphology senescence scoring. *Nat. Commun.* **12**, 257 (2021).

16. Campisi, J. CANCER: Suppressing Cancer: The Importance of Being Senescent. *Science* **309**, 886–887 (2005).

17. Collado, M. & Serrano, M. Senescence in tumours: evidence from mice and humans. *Nat. Rev. Cancer* **10**, 51–57 (2010).

18. Collado, M., Blasco, M. A. & Serrano, M. Cellular Senescence in Cancer and Aging. *Cell* **130**, 223–233 (2007).

19. Zhao, H. & Darzynkiewicz, Z. Biomarkers of Cell Senescence Assessed by Imaging Cytometry. in *Cell Senescence* (eds. Galluzzi, L., Vitale, I., Kepp, O. & Kroemer, G.) vol. 965 83–92 (Humana Press, 2013).

20. Goldman, R. D. et al. Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson–Gilford progeria syndrome. *Proc. Natl. Acad. Sci.* **101**, 8963–8968 (2004).

21. Martins, F., Sousa, J., Pereira, C. D., Cruz e Silva, O. A. B. & Rebelo, S. Nuclear envelope dysfunction and its contribution to the aging process. *Aging Cell* **19**, (2020).

22. Baus, F. Permanent cell cycle exit in G2 phase after DNA damage in normal human fibroblasts. *EMBO J.* **22**, 3992–4002 (2003).

23. Gire, V. & Dulić, V. Senescence from G2 arrest, revisited. *Cell Cycle* **14**, 297–304 (2015).
24. Kassani, S. H., Kassani, P. H., Wesolowski, M. J., Schneider, K. A. & Deters, R. Breast Cancer Diagnosis with Transfer Learning and Global Pooling. in *2019 International Conference on Information and Communication Technology Convergence (ICTC)* 519–524 (IEEE, 2019). doi:10.1109/ICTC46691.2019.8939878.

25. Tomita, H. *et al.* Deep Learning for the Preoperative Diagnosis of Metastatic Cervical Lymph Nodes on Contrast-Enhanced Computed ToMography in Patients with Oral Squamous Cell Carcinoma. *Cancers* **13**, 600 (2021).

26. Debacq-Chainiaux, F., Erusalimsky, J. D., Campisi, J. & Toussaint, O. Protocols to detect senescence-associated beta-galactosidase (SA-βgal) activity, a biomarker of senescent cells in culture and in vivo. *Nat. Protoc.* **4**, 1798–1806 (2009).

27. Gal, Y. & Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *ArXiv150602142 Cs Stat* (2016).

28. Wen, Y., Vicol, P., Ba, J., Tran, D. & Grosse, R. Flipout: Efficient Pseudo-Independent Weight Perturbations on Mini-Batches. *ArXiv180304386 Cs Stat* (2018).

29. Fort, S., Hu, H. & Lakshminarayanan, B. Deep Ensembles: A Loss Landscape Perspective. *ArXiv191202757 Cs Stat* (2020).

30. Hewitt, G. *et al.* Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence. *Nat. Commun.* **3**, 708 (2012).

31. Hernandez-Segura, A., Nehme, J. & Demaria, M. Hallmarks of Cellular Senescence. *Trends Cell Biol.* **28**, 436–453 (2018).

32. Petr, M. A., Tulika, T., Carmona-Marín, L. M. & Scheibye-Knudsen, M. Protecting the Aging Genome. *Trends Cell Biol.* **30**, 117–132 (2020).

33. Keijzers, G., Bakula, D. & Scheibye-Knudsen, M. Monogenic Diseases of DNA Repair. *N. Engl. J. Med.* **377**, 1868–1876 (2017).

34. Moreno-Blas, D. *et al.* Cortical neurons develop a senescence-like phenotype promoted by dysfunctional autophagy. *Aging* **11**, 6175–6198 (2019).
35. Wang, C. et al. DNA damage response and cellular senescence in tissues of aging mice: Senescent cells in aging mice. *Aging Cell* **8**, 311–323 (2009).

36. Burton, D. G. A. & Krizhanovsky, V. Physiological and pathological consequences of cellular senescence. *Cell. Mol. Life Sci.* **71**, 4373–4386 (2014).

37. He, S. & Sharpless, N. E. Senescence in Health and Disease. *Cell* **169**, 1000–1011 (2017).

38. Nelson, D. M., McBryan, T., Jeyapalan, J. C., Sedivy, J. M. & Adams, P. D. A comparison of oncogene-induced senescence and replicative senescence: implications for tumor suppression and aging. *AGE* **36**, 9637 (2014).

39. Idda, M. L. et al. Survey of senescent cell markers with age in human tissues. *Aging* **12**, 4052–4066 (2020).

40. Liggett, W. H. & Sidransky, D. Role of the p16 tumor suppressor gene in cancer. *J. Clin. Oncol.* **16**, 1197–1206 (1998).

41. Campisi, J., Andersen, J. K., Kapahi, P. & Melov, S. Cellular senescence: A link between cancer and age-related degenerative disease? *Semin. Cancer Biol.* S1044579X11000502 (2011) doi:10.1016/j.semcancer.2011.09.001.

42. Burd, C. E. et al. Monitoring Tumorigenesis and Senescence In Vivo with a p16INK4a-Luciferase Model. *Cell* **152**, 340–351 (2013).

43. Wang, B., Kohli, J. & Demaria, M. Senescent Cells in Cancer Therapy: Friends or Foes? *Trends Cancer* **6**, 838–857 (2020).

44. Pathak, R. U., Soujanya, M. & Mishra, R. K. Deterioration of nuclear morphology and architecture: A hallmark of senescence and aging. *Ageing Res. Rev.* **67**, 101264 (2021).

45. Filippi-Chiela, E. C. et al. Nuclear Morphometric Analysis (NMA): Screening of Senescence, Apoptosis and Nuclear Irregularities. *PLoS ONE* **7**, e42522 (2012).

46. Neri, F., Basisty, N., Desprez, P.-Y., Campisi, J. & Schilling, B. Quantitative Proteomic Analysis of the Senescence-Associated Secretory Phenotype by Data-Independent Acquisition. *Curr. Protoc.* **1**, e32 (2021).
47. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (eds. Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F.) vol. 9351 234–241 (Springer International Publishing, 2015).
Figure Legends

**Figure 1** Nuclear morphology is an accurate senescence predictor in cultured cells.  

- **a** Analysis workflow. 
- **b** Sample nuclei for controls, replicative senescence (RS) and ionizing radiation (IR) induced senescent cells. 
- **c** Area of identified nuclei (n=6,976-68,971, mean ± 95% CI). 
- **d** Convexity of identified nuclei (n=6,976-68,971, mean ± 95% CI). 
- **e** Aspect ratio of identified nuclei (n=6,976-68,971, mean ± 95% CI). 
- **f** Scatter plot of individual nuclei, with overall distributions for each at the top and right margins. 
- **g** Cell cycle analysis after exposure to several doses of IR; mn: multinucleated cells. 
- **h** Accuracy of a deep neural network (DNN) predictor on validation data. 
- **i** Receiver operating characteristics (ROC) curve of the DNN. 
- **j** Percent of nuclei in each state classified as senescent for independent cell lines. 
- **k** Distribution of prediction probabilities for several doses of IR for three fibroblast cell lines (n=38,284-106,132). 
- **l** Distribution of p21 intensities for several doses of IR for three fibroblast cell lines (n=38,284-106,132). 
- **m** Distribution of PCNA intensities for several doses of IR for three fibroblast cell lines (n=38,284-106,132). 
- **n** Correlation between predicted senescence and nearby SA-β-gal regions, showing all and 90% confidence predictions only for RS and IR groups. 
- **o** Correlation between predicted senescence and multiple markers, showing all, filtered for markers with strong signals, and filtered with 90% confidence predictions only. 
- **p** Accuracy of DNNs trained and predicting after different normalization methods. 
- **q** Correlation between morphological metrics and predicted senescence by class, BG: background.

**Figure 2** Predictions from deep ensembles increases accuracy.  

- **a** Heatmap of variation in predictions by members of ensemble (500 sample nuclei as rows, ensemble members as columns). Blue is young/control and white is senescent. 
- **b** Heatmap of per-class accuracy for control and senescent by ensemble model. 
- **c** Accuracy of deep ensemble. 
- **d** ROC curve for the
deep ensemble. e Accuracy of single model, Bayesian neural networks, deep ensemble, and bagging. f Accuracy of deep ensemble with normalized samples. g ROC curve for the deep ensemble with normalized samples. h Accuracy of three-state senescence ensemble model. i ROC curve for the three-state senescence ensemble model. j Accuracy of RS-only model. k Accuracy of IR-only model.

Figure 3 Senescence can be predicted across tissues and species. a Number of γH2AX foci by type of senescence (n=1,831-15,560, mean ± 95% CI). b Number of 53BP1 foci by type of senescence (n= 1,831-15,560, mean ± 95% CI). c Correlation between foci count and predicted senescence. d Representative immunohistochemistry micrographs of premature aging nuclei with DNA damage foci staining of γH2AX and 53BP1, HGPS: Hutchinson-Gilford Progeria Syndrome, AT: ataxia telangiectasia, CS: Cockayne Syndrome. e Nuclear area for premature aging diseases (n=4,340-15074, mean ± 95% CI), HGPS: Hutchinson-Gilford Progeria Syndrome, AT: ataxia telangiectasia, CS: Cockayne Syndrome. f Number of γH2AX foci for premature aging diseases (n= 5,162-17,584, mean ± 95% CI). g Number of 53BP1 foci by premature aging diseases (n= 5,162-17,584, mean ± 95% CI). h Predicted probability of senescence for premature aging disease (n=5,162-17,584, mean ± 95% CI). i Nuclei with nearby SA-β-gal regions. j Correlation between predicted senescence and nearby SA-β-gal and also number of samples, as confidence filtering is applied across a range of thresholds. k DAPI intensities for premature aging diseases and controls (n=4,340-15,074, mean ± 95% CI). l Representative immunohistochemistry micrographs of senescent murine astrocytes with DNA damage foci staining of γH2AX and 53BP1. m Nuclear area of murine astrocytes (n=4,888-13,549, mean ± 95% CI). n Nuclear area of murine neurons (n=33,303-62,847, mean ± 95% CI). o Number of γH2AX foci for murine astrocytes (n=4,918-13,661, mean ± 95% CI). p Number of 53BP1 foci for murine astrocytes (n= 4,918-13,661, mean ± 95% CI). q Number of γH2AX foci for murine neurons (n=33,303-62,847, mean ± 95% CI). r Number of 53BP1 foci for...
murine neurons (n=33,303-62,847, mean ± 95% CI). s Predicted senescence for murine astrocytes (n= 4,918-13,661, mean ± 95% CI). t Predicted senescence for murine neurons (n=33,303-62,847, mean ± 95% CI).

Figure 4 Senescence can be predicted across tissues and species. a Analysis workflow. b Mean nuclear area per mouse by age (n=5). c Mean nuclear convexity per mouse by age (n=5). d Mean nuclear aspect ratio per mouse by age (n=5). e Prediction probability for RS senescent (n=5). f Prediction probability for IR senescent (n=5). g Predicted percent that are RS senescent (n=5). h Predicted percent that are IR senescent (n=5). i Mean probability of predicted RS senescence for each individual by p21 state. j Mean probability of predicted IR senescence for each individual by p21 state. k PCNA intensity after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). l Area after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). m Convexity after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). n Aspect after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). o DAPI intensity after senescence induction for three fibroblast cell lines (n=30,957-119,669, mean ± 95% CI). p Predicted probability of senescence after senescence induction, using the RS model (n=30,957-119,669, mean ± 95% CI). q Accuracy of doxorubicin-only model. r Accuracy of ATV/r-only model. s Accuracy of Antimycin-A-only model. t Accuracy of unified model.

Figure 5 Nuclear morphology predict senescence and multiple diseases in humans. a Predicted probability of senescence for p21- and p21+ nuclei in human dermis, across a range of thresholds. b Analysis workflow. c Mean nuclear area per patient by age (n=148). d Mean nuclear convexity per patient by age (n=148). e Mean nuclear aspect ratio per patient by age (n=148). f Predicted percent that are RS senescent (n=169). g Predicted percent that are IR senescent (n=169). h Volcano plot of ICD-10 chapters based on IR senescence residuals and...
chi-squared p-values. \textit{i} Volcano plot of ICD-10 chapters based on RS senescence residuals and chi-squared p-values. \textit{j} Volcano plot of ICD-10 chapters based on doxorubicin senescence residuals and chi-squared p-values. \textit{k} Volcano plot of ICD-10 chapters based on ATV/r senescence residuals and chi-squared p-values. \textit{l} Volcano plot of ICD-10 chapters based on antimycin-A senescence residuals and chi-squared p-values. \textit{m} Volcano plot of ICD-10 chapters based on unified senescence residuals and chi-squared p-values. \textit{n} Table summarizing disease conditions, percent of individuals with the condition for positive and negative residual groups, p-value from chi-squared test or Fisher’s exact test, and relative risk ratio with 95% CI when significant.
Supplementary Figures
**Figure S1 Cell culture models of senescence.**

a Schematic of Ionizing radiation (IR)-induced senescence.

b Representative growth curve of cells undergoing replicative senescence (RS).

c Representative micrographs of SA-β-gal activity and DAPI in control, IR and RS cells.

d Relative p21 mRNA expression levels by qPCR (n=3, ± SEM).

e Relative p16 mRNA expression by qPCR (n=3, ± SEM).

f Relative IL6 mRNA expression by qPCR (n=3, ± SEM).

g Representative immunohistochemistry micrographs of nuclei with p21, p16, and p53 staining in control and IR cells.

h Distribution of p16 intensities for IR and control for three fibroblast cell lines (n=9,196-27,716).

i Distribution of p21 intensities for IR and control for three fibroblast cell lines (n=13,678-32,730).

j Distribution of p53 intensities for IR and control for three fibroblast cell lines (n=12,844-31,100).

k Cell count following irradiation.

l DAPI intensities for IR, RS, and control for three fibroblast cell lines (n=2,641-21,954).
Figure S2 Performance characteristics of several models.  

a Accuracy of raw model, tested on independent cell lines.  
b ROC curve of the raw model for independent cell lines.  
c Precision/Recall curve.  
d Accuracy of three-state senescent type detector model.  
e ROC curve with RS and IR for the type model.  
f Accuracy of normalized model.  
g ROC curve for the normalized model.  
h Accuracy of three-state senescent type detector model with normalized samples.  
i ROC curve for the type model with normalized samples.  
j Accuracy of Xception-based Bayesian neural network.  
k ROC curve for the Xception BNN.  
l Accuracy of InceptionV3-based Bayesian neural network.  
m ROC curve for the InceptionV3 BNN.
Figure S3 Improving performance by adjusting thresholds.  

- A Histogram of predicted probabilities for Bayesian neural network. 
- B Histogram of predicted probabilities for deep ensemble. 
- C Accuracy and percent of samples evaluated with different classifier thresholds for single model. 
- D Accuracy and percent of samples evaluated with different classifier thresholds for Xception-based Bayesian neural network. 
- E Accuracy and percent of samples evaluated with different classifier thresholds for deep ensemble with normalized samples. 
- F Accuracy and percent of samples evaluated with different classifier thresholds for ensemble of RS-only for normalized samples. 
- G Accuracy and percent of samples evaluated with different classifier thresholds for ensemble of IR-only for normalized samples.
Figure S4 Development of Senescent Phenotype. 

a Predicted probability of senescence for several time points of IR-induced senescent cells for three fibroblast cell lines (n=35,191-106,549). 

b Distribution of predicted probability split by time points. 

c Mean PCNA intensity per nucleus for several time points for three fibroblast cell lines (n=35,191-106,549). 

d Mean p21 intensity per nucleus for several time points for three fibroblast cell lines (n=35,191-106,549). 

e Mean DAPI intensity per nucleus for several time points for three fibroblast cell lines (n=35,191-106,549). 

f p21 for several time points, split by predicted senescence state. 

g PCNA for several time points, split by predicted senescence state. 

h DAPI for several time points, split by predicted senescence state. 

i Area for several time points, split by predicted senescence state. 

j Convexity for several time points, split by predicted senescence state. 

k Aspect for several time points, split by predicted senescence state. 

l Cell counts for several time points.
Figure S5 DNA Damage Foci and Morphological Characteristics of Predicted Senescence. a Representative immunohistochemistry micrographs of nuclei with DNA damage foci staining of gH2AX and 53BP1. b 2D histogram (density heatmap) of predicted senescence and foci count per senescence type. c 2D histogram (density heatmap) of predicted senescence and foci count for premature aging diseases. d 2D histogram (density heatmap) of predicted senescence and foci count for murine astrocytes.
Figure S6 Senescence metrics in human dermal fibroblasts. 

a Sample of human dermis, stained with H&E and DAB for p21. 
b Predicted probability of senescence from IR model and number of nuclei for p21- and p21+ nuclei in human dermis. 
c Predicted probability of senescence from Doxorubicin model and number of nuclei for p21- and p21+ nuclei in human dermis. 
d Predicted probability of senescence from ATV/r model and number of nuclei for p21- and p21+ nuclei in human dermis. 
e Predicted probability of senescence from antimycin-A model and number of nuclei for p21- and p21+ nuclei in human dermis. 
f Predicted probability of RS senescence (n=169). 
g Predicted probability of IR senescence (n=169). 
h Predicted probability of doxorubicin senescence (n=169). 
i Predicted probability of ATV/r senescence (n=169). 
j Predicted probability of antimycin-A senescence (n=169). 
k Percent of doxorubicin senescence (n=169). 
l Percent of ATV/r senescence (n=169). 
m Percent of antimycin-A senescence (n=169). 
n 2D histograms (density heatmaps) of nuclei from the histology slides.
Figure S7 Additional Metrics and Clinical Codes. a Correlation coefficient matrix of metrics from nuclei. b Contour distribution of aspect versus area of nuclei for individuals under 40. c Contour distribution of aspect versus area of nuclei for individuals over 60. d Volcano plot of conditions based on IR senescence residuals and Fisher's exact test p-values. e Volcano plot of conditions based on RS senescence probability residuals and Fisher's exact test p-values. f Volcano plot of conditions based on doxorubicin senescence probability residuals and Fisher's exact test p-values. g Volcano plot of conditions based on ATV/r senescence probability residuals and Fisher's exact test p-values. h Volcano plot of conditions based on antimycin-A senescence probability residuals and Fisher's exact test p-values. i Volcano plot of conditions based on percent IR senescence residuals and Fisher's exact test p-values. j Volcano plot of conditions based on percent RS senescence probability residuals and Fisher's exact test p-values. k Volcano plot of conditions based on percent doxorubicin senescence probability residuals and Fisher's exact test p-values. l Volcano plot of conditions based on percent ATV/r senescence probability residuals and Fisher's exact test p-values. m Volcano plot of conditions based on percent antimycin-A senescence probability residuals and Fisher's exact test p-values. n Volcano plot of conditions based on percent unified senescence probability residuals and Fisher's exact test p-values.
| Condition                  | IR  | RS  | Doxo | ATV/r | Anti | Uni | IR%  | RS%  | Doxo% | ATV/r% | Anti% | Uni% |
|---------------------------|-----|-----|------|-------|------|-----|------|------|-------|--------|-------|------|
| Neoplasm                  | 0.005 | 0.002 |      | 0.005 |      |     | 0.006 |      |       |        |       |      |
| Malignancy                | 0.018 | 0.058 | 0.008 |       |      |     |       |      |       |        |       |      |
| Osteoporosis (DM819)      |      |      |      |       |      |     | 0.025 |      |       |        |       |      |
| Osteoarthritis (DM170)    |      |      |      |       |      |     | 0.031 |      |       |        |       |      |
| Osteoarthritis (DM199)    |      |      |      |       |      |     | 0.037 |      |       |        |       |      |
| Hypertension (DI109)      |      |      |      |       |      |     |       |      | 0.029 |        |       |      |
| Hearing loss (DH911)      |      | 0.014 | 0.014 | 0.034 |      |     |       |      |       |        |       |      |
| Hearing loss (DH919)      |      |      |      | 0.032 |      |     |       |      |       |        |       |      |
| Cerebral Infarction (DI639) |      |      |      |       |      |     | 0.013 | 0.027 |       |        |       |      |
| Hyperlipidemia (DE785)    |      |      |      |       |      |     | 0.037 |      |       |        |       |      |
| Hypercholestereima (DE780) | 0.048 | 0.047 |      |       |      |     | 0.044 |      |       |        |       |      |

**Figure S8 Table of Significant Conditions and Additional Conditions.** a Table of selected conditions with significant p values.