Age acquired skewed X chromosome inactivation is associated with adverse health outcomes in humans

Amy L Roberts[1]*, Alessandro Morea[1,2†], Ariella Amar[2], Antonino Zito[1‡, §], Julia S El-Sayed Moustafa[1], Max Tomlinson[1,2], Ruth CE Bowyer[1], Xinyuan Zhang[1], Colette Christiansen[1], Ricardo Costeira[1], Claire J Steves[1], Massimo Mangino[1,3], Jordana T Bell[1], Chloe CY Wong[4], Timothy J Vyse[2], Kerrin S Small[1]*

[1]Department of Twin Research & Genetic Epidemiology, King's College London, London, United Kingdom; [2]Department of Medical and Molecular Genetics, King's College London, London, United Kingdom; [3]NIHR Biomedical Research Centre, Guy's and St Thomas' Foundation Trust, London, United Kingdom; [4]Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

*For correspondence:
amy.roberts@kcl.ac.uk (ALR);
kerrin.small@kcl.ac.uk (KSS)

Present address: †The FIRC Institute of Molecular Oncology, Italian Foundation for Cancer Research, Milan, Italy; ‡Department of Molecular Biology, Massachusetts General Hospital, Boston, United States; §Department of Genetics, Harvard Medical School, Boston, United States

Abstract

Background: Ageing is a heterogenous process characterised by cellular and molecular hallmarks, including changes to haematopoietic stem cells and is a primary risk factor for chronic diseases. X chromosome inactivation (XCI) randomly transcriptionally silences either the maternal or paternal X in each cell of 46, XX females to balance the gene expression with 46, XY males. Age acquired XCI-skew describes the preferential selection of cells across a tissue resulting in an imbalance of XCI, which is particularly prevalent in blood tissues of ageing females, and yet its clinical consequences are unknown.

Methods: We assayed XCI in 1575 females from the TwinsUK population cohort using DNA extracted from whole blood. We employed prospective, cross-sectional, and intra-twin study designs to characterise the relationship of XCI-skew with molecular and cellular measures of ageing, cardio-vascular disease risk, and cancer diagnosis.

Results: We demonstrate that XCI-skew is independent of traditional markers of biological ageing and is associated with a haematopoietic bias towards the myeloid lineage. Using an atherosclerotic cardiovascular disease risk score, which captures traditional risk factors, XCI-skew is associated with an increased cardiovascular disease risk both cross-sectionally and within XCI-skew discordant twin pairs. In a prospective 10 year follow-up study, XCI-skew is predictive of future cancer incidence.

Conclusions: Our study demonstrates that age acquired XCI-skew captures changes to the haematopoietic stem cell population and has clinical potential as a unique biomarker of chronic disease risk.

Introduction

Ageing is a heterogenous process characterised by cellular and molecular hallmarks and can manifest clinically as frailty and multimorbidity (*Clegg et al., 2013*; *López-Otín et al., 2013*). Ageing is a primary risk factor for diseases such as cardiovascular disease and cancer, and a better understanding of the biomarkers of ageing promises to reduce the burden of chronic disease which significantly impacts the human healthspan (*López-Otín et al., 2013*).

X chromosome inactivation (XCI) evolved in placental mammals to compensate for the X-linked gene dosage between XX females and XY males. XCI transcriptionally silences either the maternal or paternal X in each cell to equalise the gene expression between 46, XX females and 46, XY males (*Lyon, 1961*). The selection of which X is silenced is a random process that occurs during development, with the XCI status then clonally inherited by all daughter cells. Therefore, mammalian female tissues are mosaics with respect to XCI status, with an expected ratio of 1:1.

However, some individuals display a skewed pattern of XCI (XCI-skew), which is defined as a deviation from the expected 1:1 ratio. Examples of primary XCI-skew have been identified, including stochastic events resulting in XCI-skew across all tissues (*Tukiainen et al., 2016*) or the preferential selection of cells expressing functioning alleles in heterozygous females with X-linked recessive traits (*Busque and Gilliland, 1998*; *Nyhan et al., 1970*). However, secondary or age acquired XCI-skew is more common and refers to increasing XCI-skew with age, particularly in mitotically active blood tissue (*Busque et al., 1996*; *Gale et al., 1997*). Within individuals, the correlation of XCI ratios between blood and other tissues diminishes over the life course as the XCI ratios in blood continue to skew with age (*Bolduc et al., 2008*; *Zito et al., 2019*).

The stability of XCI-skew in blood has been demonstrated over 18–24 months and is thought to be a gradual process affecting the whole haematopoietic stem cell population rather than representing fluctuations in the active stem cell pool (*Tonon et al., 1998*; *van Dijk et al., 2002*). Therefore, though XCI-skew is a sex-specific measurable phenotype, it is a potential marker of stem cell depletion or polyclonal expansion of haematopoietic stem cells, which are age-associated traits irrespective of chromosomal sex (*Busque et al., 1996*; *Busque et al., 2012*; *Gale et al., 1997*). Age acquired XCI-skew has previously been linked to autoimmunity, which presents with a stark sex-imbalance (*Chabchoub et al., 2009*), as well as breast and ovarian cancers, albeit with inconsistent findings (*Kristiansen et al., 2002*; *Lose et al., 2008*; *Manoukian et al., 2013*; *Struewing et al., 2006*). Yet the consequences of XCI-skew on chronic disease risk in an unselected population have largely been unexplored.

Clonal expansion of haematopoietic stem cells is also measurable by somatic mutations shared across blood cells, indicating a common stem cell precursor (*Xie et al., 2014*). Clonal haematopoiesis of indeterminate potential (CHIP) is a cellular phenotype describing a pre-malignant state in which ≥4% of blood cells harbour the same somatic mutation (*Jaiswal and Ebert, 2019*), thus representing monoclonal expansion. CHIP is robustly associated with all-cause mortality (*Jaiswal et al., 2014*), haematological cancers (*Genovese et al., 2014*), and cardiovascular disease (*Jaiswal et al., 2017*). XCI-skew can sometimes be a marker of CHIP: XCI-skew was previously used to determine clonality (*Busque and Gilliland, 1998*), and exome sequencing of females with XCI-skew identified *TET2* mutations as drivers of pre-malignant clonal haematopoiesis. However, XCI-skew and CHIP are not completely mutually inclusive (*Busque et al., 2012*).

Given XCI-skew is potentially tagging changes to the haematopoietic stem cell pool, we hypothesised that XCI-skew may be a marker of biological ageing and a risk factor for chronic disease. We tested this hypothesis by assaying XCI-skew in 1575 females from the TwinsUK cohort and employed prospective, cross-sectional, and intra-twin study designs to characterise the relationship of XCI with molecular and cellular measures of ageing, cardiovascular disease risk, and cancer diagnoses.

Methods

TwinsUK cohort

Archival DNA samples derived from whole blood (collected 1997–2017) were selected from individuals of the TwinsUK population cohort (*Verdi et al., 2019*). Twin pairs were date matched and the final dataset of 1575 samples comprised 423 monozygotic (MZ) twin pairs ($n_{individuals}$ = 846), 257 dizygotic (DZ) twin pairs ($n_{individuals}$ = 514), and 215 singletons (*Figure 1—figure supplement 2*). The age range of the XCI cohort is 19–99, with a median age of 61 (Figure 2A). All samples and information were collected with written and signed informed consent, including consent to publish within the TwinsUK study. TwinsUK has received ethical approval associated with TwinsUK Biobank (19/NW/0187), TwinsUK (EC04/015) or Healthy Ageing Twin Study (HATS) (07 /H0802/84) studies from NHS Research Ethics Service Committees London – Westminster.

Human Androgen Receptor Assay (HUMARA)

The HUMARA method combines methylation-sensitive restriction enzyme digest and amplification of a highly polymorphic (CAG)n repeat in the first exon of the X-linked *AR* gene, allowing for the differentiation of the active and inactive chromosomes in heterozygous individuals (*Cutler Allen et al., 1992*). Here, 625 ng of genomic DNA was divided into three aliquots and incubated for 30 min at 37 °C with (i) the methylation-sensitive enzyme *HpaII*, (ii) the methylation-insensitive enzyme *MspI*, or (iii) water (mock digest) in 1 × New England Biolabs CutSmart Buffer. The *HpaII* digest was followed by an additional 20 min at 80 °C to avoid residual enzymatic activity. Fluorescently labelled PCR primers (FAM, VIC, NED, or PET; Forward primer 5'-dye-GCTGTGAAGGTTGCTGTTCCTCAT-3', Reverse primer 5'-TCCAGAATCTGTTCCAGAGCGTGC-3') were used in New England BioLabs One *Taq* Master Mix to amplify 1.5 µl of digested PCR product. The Mock and *HpaII* digested DNA were amplified in triplicate (using FAM, VIC, and NED), and the *MspI* digest, used as control of digestion efficiency, was amplified once (using PET). All PCRs were amplified with an initial denaturation step at 94 °C for 5 min, followed by 30 cycles of 94 °C for 30 s, 60 °C for 1 min, and 72 °C for 2 min, and a final elongation step of 72 °C for 15 min. To minimize technical bias and batch effects, the labelled amplified products were diluted 1:15 with nuclease-free ddH_2O and pooled together with the GeneScan 500 LIZ size standard and analysed on an ABI 3730xl. Twin pairs were assayed on the same plate and plates contained a mix of both MZ and DZ pairs. Two replicates were included on each plate and a within-plate correlation of 0.99 was measured. A total of 2382 DNA samples were assayed, with 194 failed samples, and 601 samples were homozygous for the CAG repeat and were therefore uninformative (*Figure 1—figure supplement 2*).

Calculation of XCI

Data were analysed using the Microsatellite Analysis Software available on the Thermo Fisher Cloud. The XCI status was calculated in each of the triplicates as follows:

- Allele Ratio Mock Digestion (Rm)=allele 1 peak height / allele 2 peak height
- Allele Ratio *HpaII* Digestion (Rh)=allele 1 peak height / allele 2 peak height
- Normalized Ratio (Rn)=Rh/Rm
- XCI percentage = [Rn/(Rn +1)] * 100

A coefficient of variation (CV) was calculated across the triplicates and samples with CV >0.15 were excluded from downstream analysis (n=12; *Figure 1—figure supplement 2*). A mean XCI percentage (0–100%) was calculated for each sample, where 50% is perfectly balanced XCI and the directionality of XCI away from 50% is uninformative (e.g., both 0% and 100% are considered equal). Therefore, the XCI values are collapsed to a range of 50–100% when XCI is the dependent variable in analyses.

XCI-skew categorical variable

The XCI percentage data were normalised, and a categorical XCI-skew variable was created from the absolute values of the normalised distribution as follows: standard deviation (SD)<1 from the mean = random XCI (0); 1≤SD<2 = skewed XCI (1); and SD≥2 = extreme skew (2). As such, XCI-skew equated to ≥75% XCI, and extreme XCI-skew equated to ≥91% XCI (*Figure 1—figure supplement 1*). These thresholds are very similar to previous studies (*Busque et al., 1996*; *Gale et al., 1997*), and allowed for linear associations to be tested using the XCI-skew categorical variable.

Statistical analysis

In all regression models, a linear mixed effects model was used with relatedness and family structure fitted with a random intercept using the lme4 package (*Bates et al., 2015*). The relevant fixed effects are described in each section below and were specific to each test. All analyses were carried out using R version 4.1.1, and all plots were generated using ggplot (*Wickham, 2016*).

Chronological and biological ageing

Datasets were matched to be within 1 year of the XCI DNA sample and the significance threshold after Bonferroni correction was $p<0.007$ to account for multiple testing across the seven tests in this section. Chronological age was calculated at time of DNA sampling and the association was tested using XCI as the dependent variable. Body Mass Index (BMI) measures were taken during clinical visits and obesity was defined as BMI ≥30. Smoking status was classified based on longitudinal questionnaire answers (*Christiansen et al., 2021*). The associations were tested with XCI as the dependent variable and obesity (obese/not obese) or smoking (ever/never smoker) as the independent variable, controlling for age as a fixed effect. A Frailty Index (*Searle et al., 2008*) was calculated based on longitudinal questionnaire data and used as the dependent variable, with XCI-skew as the independent variable and age and BMI as fixed effects. Leukocyte Telomere Length was measured using qPCR as previously described (*Codd et al., 2010*), and the normalised measures were used as the dependent variable and XCI-skew as the independent variable, with age and smoking as fixed effects. DNA methylation (DNAm) GrimAge was calculated using 450 K methylation data and GrimAge epigenetic age acceleration measures were obtained from regressing epigenetic age on chronological age (*Costeira et al., 2021*). GrimAge Acceleration was used as the dependent variable and XCI-skew as the independent variable, with no additional covariates included in the model.

Whole blood count data

Automated whole blood count data were date-matched to the XCI DNA sample, and each of the 10 blood count variables was normalised. The significance threshold after Bonferroni correction was $p<0.005$ to account for multiple testing across the 10 tests. In addition, Monocyte-to-Lymphocyte Ratio (MLR) and Neutrophil-to-Lymphocyte Ratio (NLR) were calculated by dividing the total monocyte or neutrophil count, respectively, by total lymphocyte count, and a Bonferroni-corrected significance threshold of $P<0.025$. Associations were tested with XCI-skew as an independent variable after controlling for age, BMI, seasonality, and smoking status as fixed effects in a linear mixed effects model.

Cytokines levels and C-reactive protein

Serum IL-1β, IL-10, IL-6, and TNF were measured simultaneously using the bead-based high sensitivity human cytokine kit (HSCYTO-60SK, Linco-Millipore) according to the manufacturers' instructions. CRP concentrations from serum were measured with the Human Cardiovascular Disease Panel 2 LINCOplex Kit (HCVD2-67BK, Linco-Millipore) and with the Extracellular Protein Buffer Reagent Kit (LHB0001, Invitrogen). CRP concentrations were diluted 1:2000 prior to analysis and assayed in duplicate, as previously described (*Ligthart et al., 2018*). Date-matched data with batch effects regressed out were normalised and associations were assessed using linear mixed models controlling for seasonality and age as fixed effects. The significance threshold after Bonferroni correction was $p<0.01$ to account for multiple testing across five tests.

Atherosclerosis and cardiovascular disease (ASCVD) risk score

The ASCVD risk score was calculated for a subset of 228 individuals with date-matched data on age, total cholesterol, HDL cholesterol, smoking, diabetes, systolic blood pressure, and hypertension medication, as previously described (*Goff et al., 2014*). A linear mixed effects model was used to control for BMI and monocyte abundance as fixed effects (*Madjid et al., 2004*). Twin pairs discordant for XCI-skew but matched for date of visit and age (n=34 pairs) were used for the intra-twin study, and ASCVD risk scores were compared using a one-sided paired samples Wilcoxon test.

Cancer and all-cause mortality

Anonymised data were obtained from the National Disease Registration Service. Study entry was the date of DNA sampling and follow-up occurred through to January 2020 (study end date). For the cancer analysis, individuals who had not experienced the event were censored at 10 years, study end date, or date-of-death. All participants with a history of cancer before sampling, or within 6 months of sampling, were excluded from analyses, and reports of non-melanoma skin cancer were filtered, leaving a sample size of 1417. For all-cause mortality, individuals who had not experienced the event were censored at 10 years or study end date. For both analyses, age, relatedness, and zygosity were controlled for in the Cox regression model using R package Survival (*Therneau, 2021*). Proportional hazards assumptions were assessed using the cox.zph function of the Survival package. Kaplan-Meier plots were used for graphical representation of years until diagnosis. XCI-skew and extreme XCI-skew groups were combined due to the limited number of events.

Results

XCI-skew in the TwinsUK population cohort

We measured XCI in DNA derived from whole blood using the methylation-sensitive PCR-based Human Androgen Receptor Assay (HUMARA), which differentiates between alleles from the active and inactive X (*Cutler Allen et al., 1992*; *Hatakeyama et al., 2004*). HUMARA is an extensively used
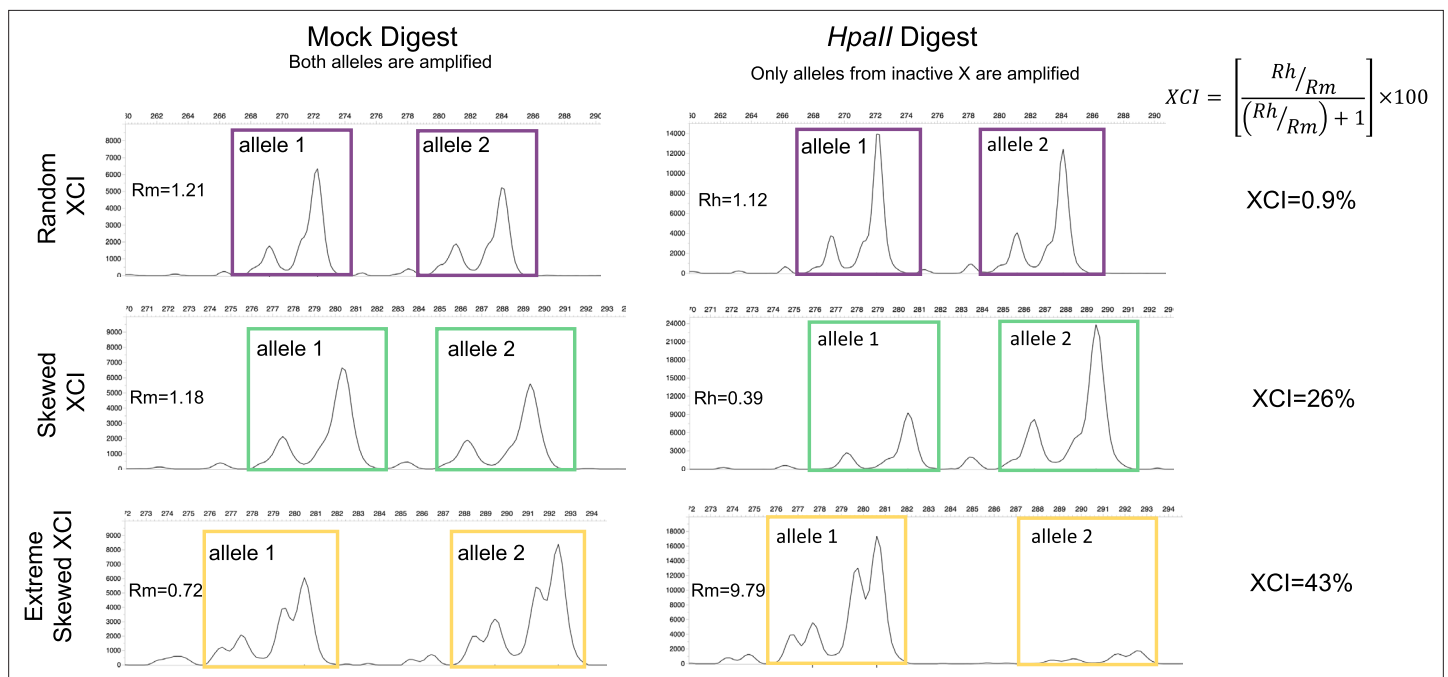

Figure 1. Measuring XCI with the HUMARA assay. The Human Androgen Receptor Assay (HUMARA) uses methylation-sensitive restriction enzyme digest and PCR to measure skewed X-inactivation. The assay estimates XCI-skew by comparing the relative abundance of allele specific fragments from a mock digest to a methylation-sensitive *HpaII* digest in which only the alleles from the inactive X are amplified. Representative examples are displayed of fragment analysis of the PCR products for samples with random XCI (top), skewed XCI (middle), and extreme skewed XCI (bottom). The x-axis shows the size, and the y-axis represents the abundance, of the PCR products, respectively. The left panel shows the PCR products after a mock digest with water, resulting in amplification of both alleles regardless of chromosomal inactivation. The right panel shows the PCR products after a restriction enzyme digest with methylation-sensitive enzyme *HpaII*, resulting in amplification of only the alleles deriving from inactive chromosomes. For each sample, the ratio of the *HpaII* digested allele products (Rh = allele 1/allele 2) is divided by the ratio of the Mock digest allele products (Rm = allele 1/allele 2) to create a Normalized Ratio (Rn). The XCI percentage is then calculated using the formula [Rn/(Rn +1)] * 100. Images were generated using the Microsatellite Analysis Software on the Thermo Fisher Cloud.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. A histogram showing the calculation of the XCI-skew categorical variable from the normalised distribution of XCI scores across the TwinsUK cohort (n=1575).

Figure supplement 2. A flowchart of sample processing and inclusion criteria for each results group.
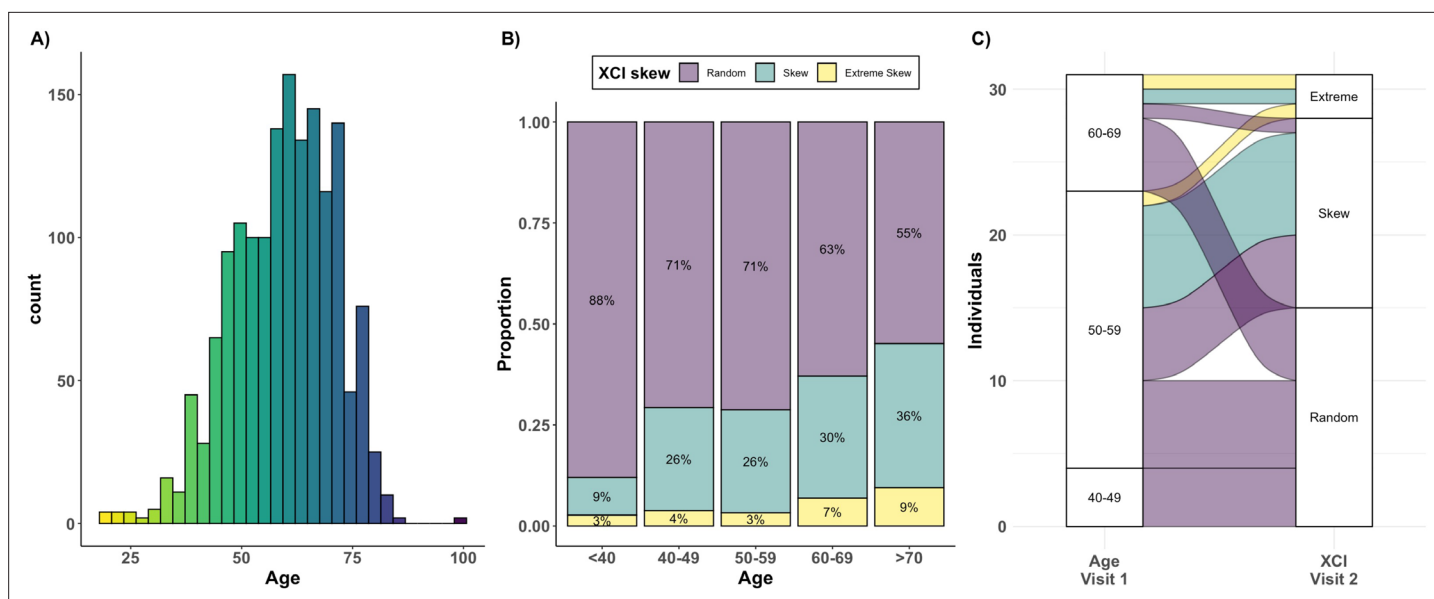
Figure 2. Age acquired XCI-skew across age groups and time. (**A**) A histogram displaying the age distribution of the TwinsUK HUMARA cohort (age range: 19–99; median age = 61). (**B**) The proportions of individuals (y-axis) in each of three XCI-skew categories across increasing age groups (x-axis) are shown (N=1575). (**C**) A Sankey plot shows the longitudinal changes to XCI in 31 individuals across two measures 15–17 years apart. Colours indicate XCI at visit 1, axis 1 displays the age group of individuals at visit 1, and axis 2 displays XCI at visit 2.

The online version of this article includes the following source data for figure 2:

Source data 1. XCI-skew across age groups.

assay which correlates well with transcription-based methods (**Bolduc et al., 2008**; **Zito et al., 2019**). The output of the HUMARA assay is a continuous XCI variable from 0–100%, where 50% is perfectly balanced XCI and the directionality of XCI away from 50% is uninformative (e.g., both 0% and 100% are considered equal). We normalised the distribution of the continuous XCI values across the cohort, and defined XCI-skew as measures 1 SD from the mean, corresponding to XCI score ≥75%, and extreme XCI-skew as measures 2 SD from the mean, corresponding to XCI score ≥91% (**Figure 1—figure supplement 1**). Of note, these values are extremely similar to thresholds previously used in the literature (**Bolduc et al., 2008**; **Busque and Gilliland, 1998**). **Figure 1** shows representative samples of how the HUMARA assay relates to the three categorical XCI-skew statuses.

We defined XCI-skew in 1575 participants (median age = 61; **Figure 2A**) unselected for chronic disease status from the TwinsUK population cohort (**Verdi et al., 2019**), which comprised of 423 MZ pairs, 257 DZ pairs, and 215 singletons. In line with previous studies (**Vickers et al., 2001**), we see increased concordance of XCI-skew within MZ twin pairs compared with DZ twin pairs: 27% of MZ twin pairs (114/423), and 45.5% of DZ twin pairs (117/257), were discordant for their categorical XCI-skew status. We date-matched the XCI data with existing phenotypes from TwinsUK (e.g., blood count data, molecular markers) in the subsets of individuals on whom each phenotype was available (**Figure 1—figure supplement 2**).

Cross-sectional and longitudinal changes to XCI-skew with age

We assessed changes in frequency of XCI-skew across increasing age groups and identified 12% (9 of 75) of individuals under 40 years old (yrs) displaying XCI-skew (≥75% XCI); 28% (183 of 652) of 40–59 yrs; 37% (185 of 498) of 60–69 yrs; and 44% (132 of 303) of those over 70 yrs (**Figure 2B**). Proportions of individuals displaying extreme XCI-skew (≥91% XCI) remains consistent at ~3–4% below the age of 60 but increases to 7% of 60–69 yrs and 9% of those over 70 yrs. These results suggest a stepwise increase in prevalence of XCI-skew happening after 40 years of age, then again after 60 years of age, where we also see the first increase in prevalence of extreme XCI-skew (**Figure 2B**). As expected, after controlling for relatedness and zygosity, we find a significant positive association between age and XCI skewing (p=2.8 × 10^{-9}, N=1575). This result replicates the many existing studies

on age acquired XCI-skew and acts as a validation of the TwinsUK HUMARA dataset (*Busque et al., 1996*; *Gale et al., 1997*; *Hatakeyama et al., 2004*; *Zito et al., 2019*).

We assessed change in XCI-skew over time using 31 individuals on whom we had an additional second sample available from 15 to 17 years prior to the main study (*Figure 2C*. median age at visit 1=55.5; median age at visit 2=72.1). The two individuals who had extreme XCI-skew at visit 1 still displayed extreme XCI-skew at visit 2. Of the eight individuals who had XCI-skew at visit 1, seven remained skewed and one progressed to extreme XCI-skew at visit 2. Of the 21 who had a random pattern of XCI at visit 1, 15 (71.4%) remained the same, and 6 (28.6%) progressed to XCI-skew at visit 2. These longitudinal data indicate that XCI-skew categorisation persists over extended periods of time and increases over the life course.

XCI-skew is independent of known markers of biological ageing

Ageing is a heterogenous process in which an individual's biological age can differ from their chronological age. Smoking and obesity are risk factors for accelerated ageing (*Tam et al., 2020*). Accelerated ageing can be estimated through measures of frailty (*Clegg et al., 2013*; *Chen et al., 2014*),
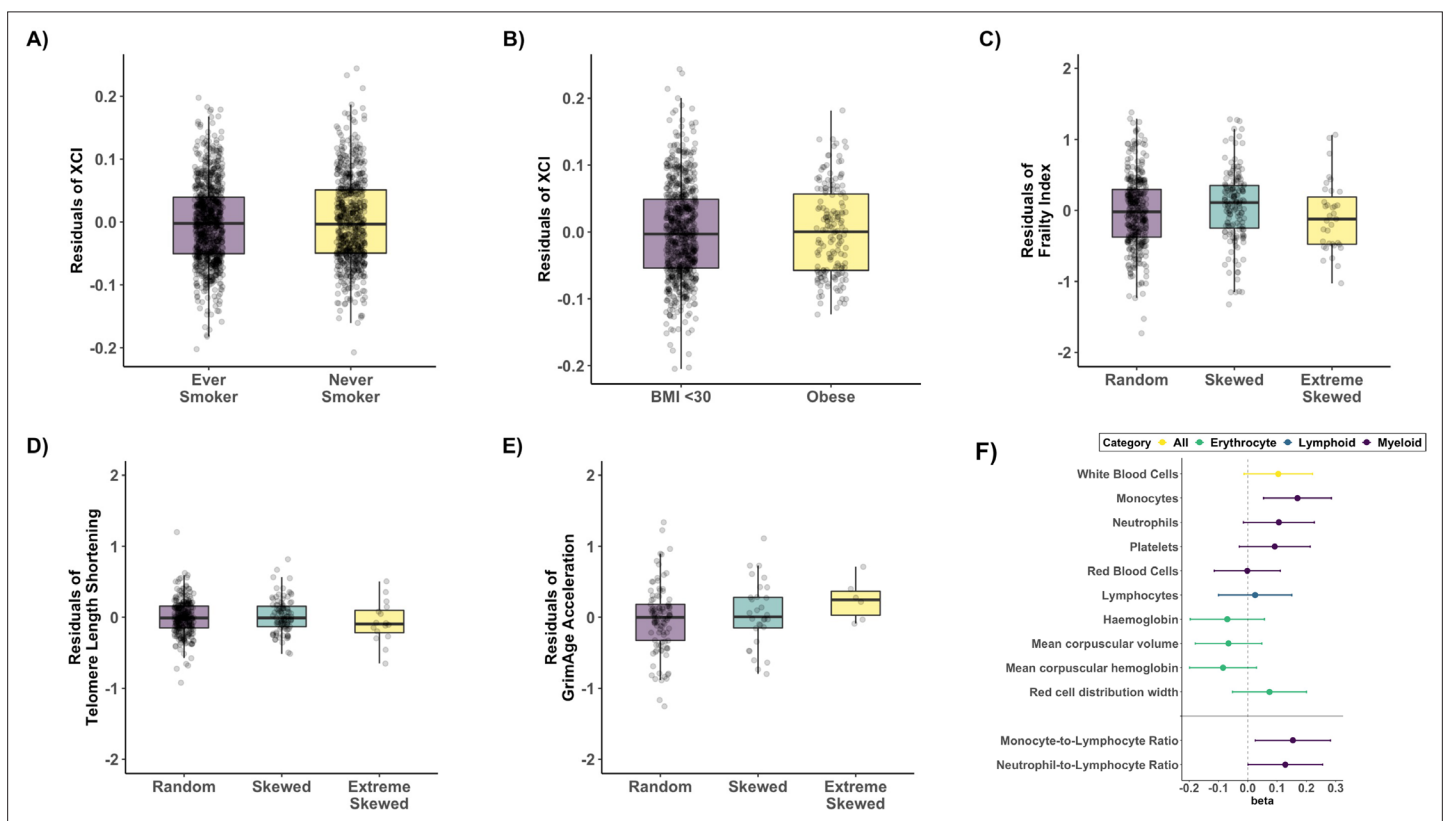

Figure 3. Age acquired XCI-skew and markers of biological ageing and blood cell counts. Box plots representing the results of the linear regression mixed effects models to assess (**A**) Smoking status (p=0.33, N_{never_smoker} = 879; N_{ever_smoker} = 673) and (**B**) obesity (p=0.88, N_{not_obese} = 726; N_{obese} = 165) after correcting for age with XCI as the dependent variable, and (**C**) frailty index (p=0.59, $N_{ranodm\ XCI}$ = 398; $n_{Skewed\ XCI}$ = 177; $N_{extreme\ skew}$ = 36), after correcting for age and BMI, (**D**) Leukocyte telomere length shortening (p=0.9, $N_{ranodm\ XCI}$ = 278; $n_{Skewed\ XCI}$ = 103; $N_{extreme\ skew}$ = 16) after correcting for age and smoking status, and (**E**) DNAm GrimAge acceleration (p=0.22, $N_{ranodm\ XCI}$ = 101; $n_{Skewed\ XCI}$ = 30; $N_{extreme\ skew}$ = 6), with XCI-skew as the dependent variable. All boxplots display the median and IQR, and have the residuals of the models on the y-axis. (**F**) A forest plot of associations with data-matched Complete Blood Count data (top panel) and Myeloid-to-lymphoid ratios (bottom) with effect size and lower and upper confidence intervals indicated. Associations were tested with XCI-skew as an independent variable after controlling for age, BMI, seasonality, and smoking status as fixed effects in a linear mixed effects model ($N_{ranodm\ XCI}$ = 445; $n_{Skewed\ XCI}$ = 183; $N_{extreme\ skew}$ = 43). The significance threshold after Bonferroni correction was p<0.005 and p<0.023 to account for multiple testing across the 10 tests and 2 tests, respectively.

The online version of this article includes the following source data and figure supplement(s) for figure 3:

Source data 1. XCI-skew and DNAm GrimAge Acceleration.

Figure supplement 1. Age acquired XCI-skew and Cytokine and CRP measures.

and on a molecular level using measures of leukocyte telomere length shortening (*Blackburn et al., 2006*) and epigenetic ageing clocks such as DNA methylation (DNAm) GrimAge (*Lu et al., 2019*); all these measures are associated with adverse health outcomes (*Blackburn et al., 2006*; *Hewitt et al., 2020*; *Lu et al., 2019*).

Given the robust association with chronological age, we sought to establish whether XCI-skew was associated with biological ageing using measures taken within 1 year of the XCI DNA sample and using linear regression mixed effects models, controlling for relatedness and zygosity as random effects (*Figure 3*). We observed no association with smoking status (p=0.33, N_{never_smoker}=879; N_{ever_smoker} = 673; *Figure 3A*), nor obesity (p=0.88, N_{not_obese}=726; N_{obese} = 165; *Figure 3B*), after correcting for age. We also found no association with a robust frailty index (p=0.59, $N_{ranodm\ XCI}$=398; $n_{Skewed\ XCI}$ = 177; $N_{extreme\ skew}$ = 36; *Figure 3C*) after correcting for age and BMI, nor with leukocyte telomere length shortening (p=0.9, $N_{ranodm\ XCI}$=278; $n_{Skewed\ XCI}$ = 103; $N_{extreme\ skew}$ = 16; *Figure 3D*), after correcting for age and smoking status. Finally, we see no association with DNAm GrimAge acceleration (p=0.22, $N_{ranodm\ XCI}$=101; $n_{Skewed\ XCI}$ = 30; $N_{extreme\ skew}$ = 6; *Figure 3E*), however, we believe the relationship between XCI-skew and epigenetic ageing would benefit from a follow-up study with a larger sample size particularly given the limited number of samples with extreme XCI-skew in our analysis. Together, these data, ranging from the molecular to organismal level, suggest age acquired XCI-skew is independent of many known markers of biological ageing and is potentially a unique biomarker with unexplored utility.

XCI-skew is associated with increased monocyte abundance and decreased IL-10 levels

Changes in blood cell composition can be indicative of ill health or systemic inflammation (*Kabat et al., 2017*; *Madjid et al., 2004*; *Patel et al., 2009*), and a haematopoietic stem cell bias towards the myeloid lineage is observed with ageing (*Pang et al., 2011*). We tested for associations between XCI-skew and whole blood count data, including white cell differentials, in a subset of individuals with matched data ($N_{ranodm\ XCI}$ = 445; $n_{Skewed\ XCI}$ = 183; $N_{extreme\ skew}$ = 43, median age = 63). After controlling for age, seasonality, BMI, smoking, relatedness and zygosity in a linear regression mixed effects model, XCI-skew is association with increased monocyte abundance after multiple testing correction (p=0.0038), and we observed nominal increases in abundance across other myeloid cells (*Figure 3F*). We next tested the hypothesis that XCI-skew was associated with a myeloid lineage bias using the Monocyte-to-Lymphocyte Ratio (MLR) (*Chen et al., 2019*) and Neutrophil-to-Lymphocyte Ratio (NLR) (*Arbel et al., 2012*), and detect an association between XCI-skew and MLR (p=0.019), and a nominal association with NLR (p=0.042) (*Figure 3F*). Though myeloid cells show a greater degree of skewing, thought to be due to the shorter lifespan of these cells (*Gale et al., 1997*), we do not believe the association seen here between XCI-skew and increasing monocyte numbers is causal: XCI-skew is defined as ≥25% shift in cell mosaicism, whereas monocytes account for only ~10% of white blood cells.

'Inflammageing' is the chronic pro-inflammatory phenotype observed in ageing and is considered an altered state of intercellular communication (*López-Otín et al., 2013*). Markers of inflammageing include cytokines produced by immune cells and C-reactive protein (CRP) produced by liver cells (*Salminen et al., 2012*). We date-matched the XCI data with serum levels of CRP ($N_{ranodm\ XCI}$ = 121; $n_{Skewed\ XCI}$ = 38; $N_{extreme\ skew}$ = 6) and a more modest cytokine dataset of interleukin (IL)–6, IL-1B, IL-10, and TNF ($N_{ranodm\ XCI}$ = 23; $n_{Skewed\ XCI}$ = 4) and used linear regression mixed effects models to control for age, seasonality, relatedness and zygosity (*Figure 3—figure supplement 1*). We see no association with primary markers of inflammageing CRP (p=0.41), IL-6 (p=0.41), or TNF (p=0.61), though a nominal association with IL-1β is observed (p=0.02) which does not pass multiple correction, but warrants follow up analysis with a larger sample size. However, we observe a strong negative association with IL-10 (p=0.0008, *Figure 3—figure supplement 1*). IL-10 is a broadly expressed anti-inflammatory cytokine which can inhibit the proinflammatory responses of both innate and adaptive immune cells (*Saraiva and O'Garra, 2010*). Though we note that due to minimal overlap in the datasets, we were unable to control for cell type composition, which may impact these findings.
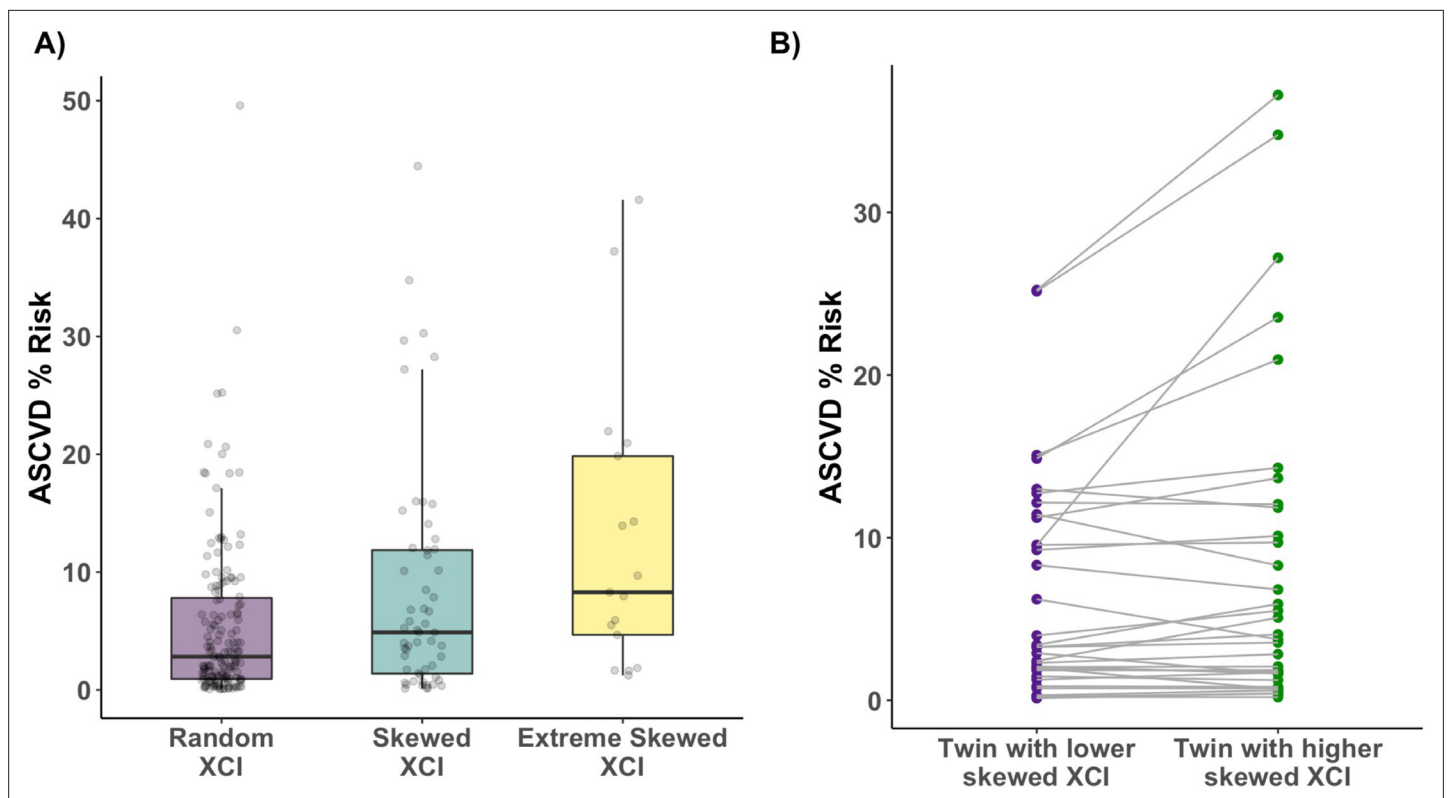
Figure 4. Age acquired XCI-skew and cardiovascular disease risk score. (**A**) Using a linear regression mixed effects model to control for BMI, monocyte count, relatedness and zygosity, XCI-skew is associated with increased atherosclerotic cardiovascular disease risk score ($N_{ranodm\ XCI}$ = 155; $n_{Skewed\ XCI}$ = 56; $N_{extreme\ skew}$ = 17; p=0.01). The boxplot displays the median and IQR. (**B**) Using age-matched twin pairs discordant for their XCI-skew status (N_{pairs} = 34), XCI-skew is associated with increased atherosclerotic cardiovascular disease risk score in the intra-twin analysis (one-sided paired samples Wilcoxon test; p=0.025).

The online version of this article includes the following source data for figure 4:

Source data 1. XCI-skew and ASCVD Risk Score.

Atherosclerotic cardiovascular disease risk is increased in individuals with XCI-skew

Cardiovascular disease (CVD) is the leading cause of death worldwide, and monocytes are an innate immune cell type known to be mediators in CVD disease progression and are found in atherosclerotic lesions (*Libby et al., 2016*). Given the association of XCI-skew with increased monocyte abundance, and the association of CHIP with CVD, we hypothesised that XCI-skew could also be associated with CVD. The atherosclerotic cardiovascular disease (ASCVD) risk score (*Goff et al., 2014*) (see Methods) captures traditional risk factors and gives a predicted risk of a major CVD event in the next 10 years, with an ASCVD risk score >7.5% representing intermediate risk, and an ASCVD risk score >20% representing high risk. In a cross-sectional study of 228 individuals ($N_{ranodm\ XCI}$ = 155; $n_{Skewed\ XCI}$ = 56; $N_{extreme\ skew}$ = 17; median age = 62) with matched health data available, XCI-skew was associated with increased ASCVD risk score after controlling for BMI, monocyte abundance, relatedness and zygosity using a linear regression mixed effects model (p=0.01, *Figure 4A*). 23.5% (4 of 17) and 35.3% (6 of 17) of individuals with extreme XCI-skew have high and intermediate ASCVD risk, respectively, compared to 4.5% (7 of 155) and 21.9% (34 of 155) of individuals with random XCI.

To ensure the observed association wasn't spuriously driven by the age component of the ASCVD risk score (see Methods), we took age matched MZ and DZ twin pairs (n_{pairs} = 34) discordant for their XCI-skew status and tested intra-twin ASCVD risk scores. The intra-twin analysis validated the association between higher XCI-skew and increased ASCVD risk score (one-sided paired samples Wilcoxon test p=0.025; *Figure 4B*), adding further support to the finding.
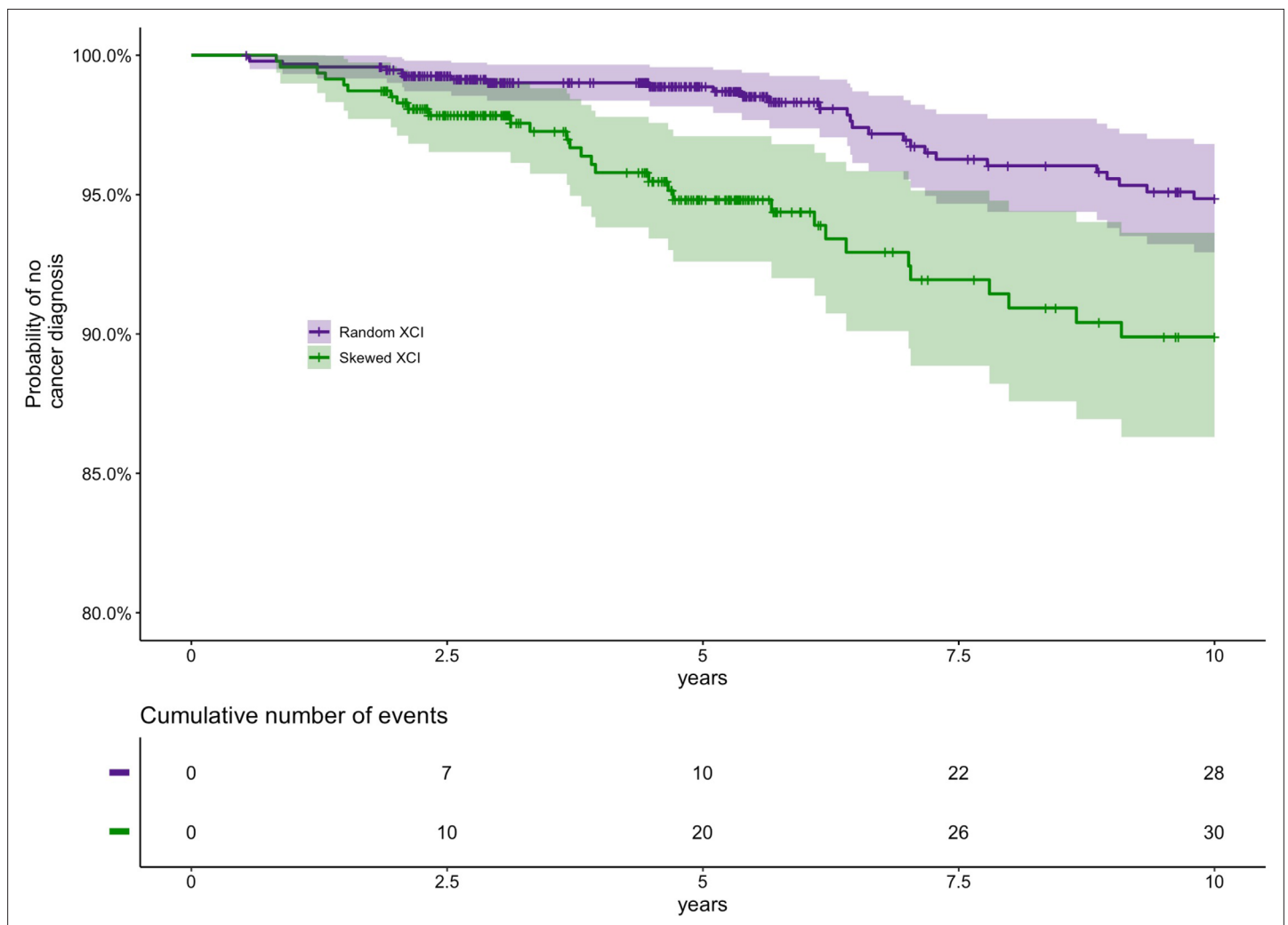
Figure 5. Prospective study of XCI-skew and future cancer diagnosis. Using a Cox regression analysis and controlling for age, relatedness and zygosity, XCI-skew is predictive of future cancer incidence in a 10 year follow-up (p=0.012; HR = 1.95). Each cancer event represents the first cancer diagnosis (excluding non-melanoma skin cancer) in each individual. A Kaplan-Meier plot (top) and cumulative events (bottom) of cancer diagnosis in individuals with XCI-skew (N=469) and random XCI (N=948) are shown. 2.9% (28/948) of individuals with random XCI, and 6.4% (30/469) of individuals with XCI-skew, developed cancer in the 10 years follow-up.

XCI-skew is predictive of future cancer diagnosis in 10-year follow-up

The association of XCI-skew with cancer has largely been assessed in case-control studies focusing on cancers of female reproductive organs, with limited replicated findings (*Buller et al., 1999*; *Kristiansen et al., 2002*; *Struewing et al., 2006*). We conducted a prospective 10 year follow-up study (median follow-up time 5.65 years) from time of DNA sampling in 1417 individuals (N$_{random}$ = 948, N$_{skewed}$ = 469, median age = 60) who were cancer-free at baseline to assess the association between XCI-skew and future cancer diagnoses (cancer events = 58; *Supplementary file 1*). Each cancer event represents the first cancer diagnosis (excluding non-melanoma skin cancer) in each individual and subsequent diagnoses are not included. Using multivariate Cox regression analysis controlling for age, relatedness and zygosity, XCI-skew was associated with increased probability of cancer diagnosis (p=0.012; Hazard Ratio (HR)=1.95 (95% Confidence Interval (CI)=1.16–3.28); *Figure 5*). Though we were underpowered to run associations for specific cancers, we were interested to explore the potential that blood cancers were partly driving the association given the robust association of CHIP with haematological cancers. However, running the analysis without the Haematopoietic/Lymphoid Tissue cancers (see *Supplementary file 1*) strengthened the observed association (p=0.009; HR = 2.04 [1.20–3.49]) suggesting a crucial difference in cancer risk between XCI-skew and CHIP.

It has been suggested that XCI-skew could have a protective effect against all-cause mortality in cohorts selected for longevity (*Mengel-From et al., 2012*). However, in our 10 year follow-up study in a population cohort (median follow-up time = 5.71 years; deaths = 41), we see a non-statistically significant trend towards a positive association with all-cause mortality (p=0.29; HR = 1.39 [0.75–2.58]).

Discussion

The age association of XCI-skew in blood tissue has long been established, with increased prevalence of females displaying XCI-skew after middle age, which is a critical time for the development of chronic disease (*Busque et al., 1996*; *Gale et al., 1997*; *Zito et al., 2019*). Although XCI-skew is measurable in a sex-specific manner, age acquired XCI-skew is potentially a marker of stem cell depletion or clonal expansion of haematopoietic stem cells (*Busque et al., 1996*; *Busque et al., 2012*; *Gale et al., 1997*; *van Dijk et al., 2002*), and could therefore have a role in age-related chronic disease, which has been robustly established for CHIP (*Jaiswal and Ebert, 2019*). In our cross-sectional study of 1575 females, we identify associations of blood-derived age acquired XCI-skew with an increased atherosclerotic cardiovascular disease risk score and increased probability of future cancer diagnosis.

Intriguingly, XCI-skew appears independent of other known markers of biological ageing, including leukocyte telomere length shortening, inflammageing, and a robust measure of frailty, making XCI-skew a unique biomarker with clinical potential. We note that though we also see no association with the mDNA GrimAge measure of accelerated epigenetic ageing, our sample size was limited, and this perhaps warrants a follow-up study to better establish the lack of correlation. Importantly, we also demonstrate that smoking and obesity are not risk factors for age acquired XCI-skew, and with a limited longitudinal dataset, that XCI-skew persists and increases over extended periods of time (median 16 years).

We find an association with increased probability of future cancer diagnosis in a 10 year follow-up study. Previous studies assessing the role of XCI-skew in cancer have typically used case-control studies of breast and ovarian cancers and have not been robustly replicated. The lack of replication is potentially explained by the heterogeneity in study design, including age-of-onset and therapy timing (*Kristiansen et al., 2002*; *Struewing et al., 2006*), *BRCA1* mutation stratification (*Kristiansen et al., 2005*; *Lose et al., 2008*; *Manoukian et al., 2013*), and threshold level for definition of XCI-skew (*Buller et al., 1999*). We controlled for potential confounding effects of cancer treatment on blood cells by excluding all individuals with a cancer diagnosis prior to study entry. Due to limited cancer events, our Cox regression analysis combined all individuals with XCI-skew (≥75% XCI) and extreme XCI-skew (≥91% XCI), which suggests even modest levels of skewing represents a greater risk of cancer.

Though CHIP is also predictive of future cancer diagnosis, the association is limited to haematological cancers (*Desai et al., 2018*). Whereas here we see blood-derived measures of XCI-skew is predictive of all future cancer diagnoses, even when the haematological cancers are removed from the analysis. Follow-up studies to assess the risk of cancer of specific tissues in individuals with XCI-skew are needed, but we hypothesise that the relationship between XCI-skew measures in blood tissue and cancer will not be causal. Instead XCI-skew is likely a marker of chronic inflammation, which can predispose to the development of cancer through increased mutagenesis and can promote tumorigenesis by shaping the tumour microenvironment to stimulate tumour growth (*Greten and Grivennikov, 2019*). However, it is interesting that common environmental factors that induce chronic inflammation, such as smoking and obesity, have not been found to be risk factors for XCI-skew in our study.

We also present an association with increased atherosclerotic cardiovascular disease risk score, which captures traditional risk factors (age, total cholesterol, HDL cholesterol, smoking, diabetes, systolic blood pressure, and hypertension medication) and estimates the risk of developing CVD in the next 10 years (*Goff et al., 2014*). We were also able to utilise the powerful discordant twin design to validate this finding, thus excluding the possibility that the cross-sectional association was driven by the age component of the ASCVD risk score. Following our finding of a correlation between blood-derived XCI-skew and risk of CVD, ascertaining whether XCI-skew is associated with incident CVD is of upmost importance. CHIP has been robustly associated with incident CVD and this association is thought to be causal (*Jaiswal et al., 2017*). Furthermore, the association of CHIP and all-cause mortality is partly driven by CVD (*Jaiswal et al., 2014*). We lacked statistical power to demonstrate an association with all-cause mortality, though the trend was towards greater risk (HR = 1.39 [0.75–2.58]).

A higher-powered study able to focus on specific causes of mortality will be of great interest and importance.

Given the existing links between CHIP and XCI-skew as two age acquired blood traits, and that CHIP mutations can be found in individuals with XCI-skew (*Busque et al., 2012*), could the association between XCI-skew and CVD risk be partly driven by CHIP? It is expected that some of the individuals with XCI-skew in our study will also harbour CHIP mutations. However, a better understanding of the co-occurrence of XCI-skew and CHIP within individuals, and the mutational burden within individuals with and without XCI-skew, is an important area of future work. Whether the co-occurrence of XCI-skew and CHIP represents an amplified risk of disease, or helps better define risk categories, will be of clinical significance to establish. As the inherited genetic risk of CHIP is fast becoming better understood, deciphering the genetic predisposition to XCI-skew will also enable the assessment of potential shared genetic susceptibility between the two traits (*Kar et al., 2022*). However, many of the negative results in our study are crucial findings given they expose differences between the risk factors of XCI-skew and CHIP. Notably, recent work has demonstrated smoking and telomere length are causal risk factors for CHIP (*Kar et al., 2022*), whereas we see no association with either trait in our study, suggesting CHIP and XCI-skew have some distinct aetiologies.

On a cellular level, we observe that increased abundance of monocytes correlates with increased XCI-skew, which is of particular interest given that monocytes/macrophages are involved in the inflammatory pathophysiology of CVD (*Libby et al., 2016*). It is important to note however that changes in monocyte counts alone do not explain the observed levels of XCI-skew. Monocytes account for ~10% of white blood cells whereas XCI-skew is defined here as ≥25% shift in cell mosaicism. Furthermore, age acquired XCI-skew has previously been shown across isolated neutrophils, monocytes, and T cells, with correlations between these fractions (*Tonon et al., 1998*; *van Dijk et al., 2002*), albeit with lower levels of skewing in lymphoid cells (*Gale et al., 1997*). Instead, XCI-skew is likely associated with the age-related haematopoietic bias toward the myeloid lineage (*Pang et al., 2011*), as we see nominal associations across other myeloid cell types in addition to the monocyte- and neutrophil-to-lymphocyte ratios. Our study also demonstrates an association of reduced levels of IL-10 with increased XCI-skew. IL-10 is an anti-inflammatory cytokine produced by a broad range of immune cells (*Saraiva and O'Garra, 2010*) and subsets of monocytes differ in their capacities to secrete IL-10 (*Skrzeczyńska-Moncznik et al., 2008*). Follow-up work on inflammatory profiles linked to XCI-skew may reveal mechanistic insights.

We derived our threshold for XCI-skew from the normalised distribution of the continuous XCI values across the cohort, and defined XCI-skew as measures 1 SD from the mean, corresponding to XCI score ≥75%, and extreme XCI-skew as measures 2 SD from the mean, corresponding to XCI score ≥91%. These values are very similar to thresholds used in the literature (*Busque et al., 1996*; *Gale et al., 1997*; *Zito et al., 2019*), but allow us to test for linear associations across the increasing thresholds of XCI-skew and demonstrate that individuals with lower levels of XCI-skew, which affects >37% of females over 60, are still at elevated risk of ASCVD and future cancer diagnoses. However, to avoid Type I error inflation, we did not run additional tests to compare the extreme XCI-skew group with the random XCI group.

There are some limitations to our study. Despite the high sample size of individuals with measured XCI, analyses are carried out in subsets of these individuals where date-matched phenotype data were available. In particular, the cytokine analyses have a low sample size and though we detect significant associations with IL-10, we may be underpowered to rule out the possibility of detecting a weaker association with other cytokines, particularly IL-1β. We were also unable to control for cell type composition in the cytokine analyses due to minimal overlap with the date-matched whole blood count data, which may have a significant impact on these findings. Though most of the measures of biological ageing were well-powered, the mDNA GrimAge analysis may benefit from replication in a larger sample size to more robustly establish the relationship between XCI-skew and accelerated epigenetic ageing. Also, with only 41 deaths across the cohort, our all-cause mortality analysis is underpowered, and a study able to focus on specific causes of mortality is warranted.

In summary, we demonstrate XCI-skew is a highly prevalent cellular phenotype in females and is associated with elevated cardiovascular disease risk and predictive of future cancer incidence. Further investigations are needed to translate the biological value of XCI-skew into clinical applications for studying the association of age-related haematopoietic changes and chronic disease associations,

regardless of chromosomal sex. Understanding the mechanisms underlying this phenomenon, whether XCI-skew is reflective of other ageing markers that increase disease risk, and whether it is therapeutically actionable, are areas of particular interest.

Acknowledgements

Additional information

Competing interests

Claire J Steves: Consultant for Zoe Ltd. The other authors declare that no competing interests exist.

Funding

Author contributions

Amy L Roberts, Data curation, Formal analysis, Supervision, Investigation, Visualization, Methodology, Writing - original draft, Project administration; Alessandro Morea, Data curation, Investigation, Writing – review and editing; Ariella Amar, Investigation; Antonino Zito, Julia S El-Sayed Moustafa, Xinyuan Zhang, Data curation, Writing – review and editing; Max Tomlinson, Ruth CE Bowyer, Colette Christiansen, Ricardo Costeira, Data curation, Formal analysis, Writing – review and editing; Claire J Steves, Jordana T Bell, Resources, Writing – review and editing; Massimo Mangino, Resources, Data curation, Writing – review and editing; Chloe CY Wong, Supervision, Methodology, Writing – review and editing; Timothy J Vyse, Resources, Supervision, Writing – review and editing; Kerrin S Small, Conceptualization, Resources, Supervision, Funding acquisition, Project administration, Writing – review and editing

Author ORCIDs

Amy L Roberts http://orcid.org/0000-0002-5704-9249

Alessandro Morea http://orcid.org/0000-0002-7919-1932
Antonino Zito http://orcid.org/0000-0003-1931-984X
Claire J Steves http://orcid.org/0000-0002-4910-0489
Massimo Mangino http://orcid.org/0000-0002-2167-7470
Timothy J Vyse http://orcid.org/0000-0003-1123-1464
Kerrin S Small http://orcid.org/0000-0003-4566-0005

Ethics

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.78263.sa1
Author response https://doi.org/10.7554/eLife.78263.sa2

Additional files

Supplementary files
- Transparent reporting form
- Supplementary file 1. Number of cancer diagnoses recorded in 10-year follow-up by organ/site.
- Reporting standard 1. STROBE statement.

Data availability

Source data are provided for Figure 2B, Figure 3E and Figure 4A. All data in the manuscript have been deposited to the TwinsUK BioResource data management team and are available by application to the Twin Research Executive Access committee (TREC) at King's College London. The TwinsUK BioResource is managed by TREC, which provides governance of access to TwinsUK data and samples. This excludes the National Disease Registration Service data which are only available to internal departmental members who are ONS accredited due to the terms of the data linkage. TwinsUK data users are bound by data sharing agreement set out in the data access application form (https://twinsuk.ac.uk/wp-content/uploads/2018/11/DTR_DataAccess_Policy_0318.pdf). This includes responsibilities with respect to third party data sharing and maintaining participant privacy. Further responsibilities include a responsibility to acknowledge data sharing.

References

Arbel Y, Finkelstein A, Halkin A, Birati EY, Revivo M, Zuzut M, Shevach A, Berliner S, Herz I, Keren G, Banai S. 2012. Neutrophil/lymphocyte ratio is related to the severity of coronary artery disease and clinical outcome in patients undergoing angiography. *Atherosclerosis* **225**:456–460. DOI: https://doi.org/10.1016/j.atherosclerosis.2012.09.009, PMID: 23040448

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**. jss.v067.i01 DOI: https://doi.org/10.18637/jss.v067.i01

Blackburn EH, Greider CW, Szostak JW. 2006. Telomeres and telomerase: the path from maize, tetrahymena and yeast to human cancer and aging. *Nature Medicine* **12**:1133–1138. DOI: https://doi.org/10.1038/nm1006-1133, PMID: 17024208

Bolduc V, Chagnon P, Provost S, Dubé MP, Belisle C, Gingras M, Mollica L, Busque L. 2008. No evidence that skewing of X chromosome inactivation patterns is transmitted to offspring in humans. *The Journal of Clinical Investigation* **118**:333–341. DOI: https://doi.org/10.1172/JCI33166, PMID: 18097474

Buller RE, Sood AK, Lallas T, Buekers T, Skilling JS. 1999. Association between nonrandom X-chromosome inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer. *Journal of the National Cancer Institute* **91**:339–346. DOI: https://doi.org/10.1093/jnci/91.4.339, PMID: 10050867

Busque L, Mio R, Mattioli J, Brais E, Blais N, Lalonde Y, Maragh M, Gilliland DG. 1996. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* **88**:59–65. DOI: https://doi.org/10.1182/blood.V88.1.59.59, PMID: 8704202

Busque L, Gilliland DG. 1998. X-inactivation analysis in the 1990s: promise and potential problems. *Leukemia* **12**:128–135. DOI: https://doi.org/10.1038/sj.leu.2400936, PMID: 9519772

Busque L, Patel JP, Figueroa ME, Vasanthakumar A, Provost S, Hamilou Z, Mollica L, Li J, Viale A, Heguy A, Hassimi M, Socci N, Bhatt PK, Gonen M, Mason CE, Melnick A, Godley LA, Brennan CW, Abdel-Wahab O,

Levine RL. 2012. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nature Genetics* **44**:1179–1181. DOI: https://doi.org/10.1038/ng.2413, PMID: 23001125

Chabchoub G, Uz E, Maalej A, Mustafa CA, Rebai A, Mnif M, Bahloul Z, Farid NR, Ozcelik T, Ayadi H. 2009. Analysis of skewed X-chromosome inactivation in females with rheumatoid arthritis and autoimmune thyroid diseases. *Arthritis Research & Therapy* **11**:R106. DOI: https://doi.org/10.1186/ar2759, PMID: 19589151

Chen X, Mao G, Leng SX. 2014. Frailty syndrome: an overview. *Clinical Interventions in Aging* **9**:433–441. DOI: https://doi.org/10.2147/CIA.S45300, PMID: 24672230

Chen H, Li M, Liu L, Dang X, Zhu D, Tian G. 2019. Monocyte/lymphocyte ratio is related to the severity of coronary artery disease and clinical outcome in patients with non-ST-elevation myocardial infarction. *Medicine* **98**:e16267. DOI: https://doi.org/10.1097/MD.0000000000016267, PMID: 31261596

Christiansen C, Castillo-Fernandez JE, Domingo-Relloso A, Zhao W, El-Sayed Moustafa JS, Tsai PC, Maddock J, Haack K, Cole SA, Kardia SLR, Molokhia M, Suderman M, Power C, Relton C, Wong A, Kuh D, Goodman A, Small KS, Smith JA, Tellez-Plaza M, et al. 2021. Novel DNA methylation signatures of tobacco smoking with trans-ethnic effects. *Clinical Epigenetics* **13**:36. DOI: https://doi.org/10.1186/s13148-021-01018-4, PMID: 33593402

Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. 2013. Frailty in elderly people. *Lancet* **381**:752–762. DOI: https://doi.org/10.1016/S0140-6736(12)62167-9, PMID: 23395245

Codd V, Mangino M, van der Harst P, Braund PS, Kaiser M, Beveridge AJ, Rafelt S, Moore J, Nelson C, Soranzo N, Zhai G, Valdes AM, Blackburn H, Mateo Leach I, de Boer RA, Kimura M, Aviv A, Goodall AH, Ouwehand W, van Veldhuisen DJ, et al. 2010. Common variants near TERC are associated with mean telomere length. *Nature Genetics* **42**:197–199. DOI: https://doi.org/10.1038/ng.532, PMID: 20139977

Costeira R, Lee KA, Murray B, Christiansen C, Castillo-Fernandez J, Ni Lochlainn M, Capdevila Pujol J, Macfarlane H, Kenny LC, Buchan I, Wolf J, Rymer J, Ourselin S, Steves CJ, Spector TD, Newson LR, Bell JT. 2021. Estrogen and COVID-19 symptoms: associations in women from the COVID symptom study. *PLOS ONE* **16**:e0257051. DOI: https://doi.org/10.1371/journal.pone.0257051, PMID: 34506535

Cutler Allen R, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. 1992. Methylation of hpaii and hhai sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *American Journal of Human Genetics* **51**:1229–1239 PMID: 1281384.

Desai P, Mencia-Trinchant N, Savenkov O, Simon MS, Cheang G, Lee S, Samuel M, Ritchie EK, Guzman ML, Ballman KV, Roboz GJ, Hassane DC. 2018. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nature Medicine* **24**:1015–1023. DOI: https://doi.org/10.1038/s41591-018-0081-z, PMID: 29988143

Gale RE, Fielding AK, Harrison CN, Linch DC. 1997. Acquired skewing of X-chromosome inactivation patterns in myeloid cells of the elderly suggests stochastic clonal loss with age. *British Journal of Haematology* **98**:512–519. DOI: https://doi.org/10.1046/j.1365-2141.1997.2573078.x, PMID: 9332303

Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, Chambert K, Mick E, Neale BM, Fromer M, Purcell SM, Svantesson O, Landén M, Höglund M, Lehmann S, Gabriel SB, Moran JL, Lander ES, Sullivan PF, Sklar P, et al. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *The New England Journal of Medicine* **371**:2477–2487. DOI: https://doi.org/10.1056/NEJMoa1409405, PMID: 25426838

Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, Robinson JG, Schwartz JS, Shero ST, Smith SC, Sorlie P, Stone NJ, Wilson PWF, Jordan HS, Nevo L, Wnek J, et al. 2014. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the american college of cardiology/american heart association task force on practice guidelines. *Circulation* **129**:S49–S73. DOI: https://doi.org/10.1161/01.cir.0000437741.48606.98, PMID: 24222018

Greten FR, Grivennikov SI. 2019. Inflammation and cancer: triggers, mechanisms, and consequences. *Immunity* **51**:27–41. DOI: https://doi.org/10.1016/j.immuni.2019.06.025, PMID: 31315034

Hatakeyama C, Anderson CL, Beever CL, Peñaherrera MS, Brown CJ, Robinson WP. 2004. The dynamics of X-inactivation skewing as women age. *Clinical Genetics* **66**:327–332. DOI: https://doi.org/10.1111/j.1399-0004.2004.00310.x, PMID: 15355435

Hewitt J, Carter B, Vilches-Moraga A, Quinn TJ, Braude P, Verduri A, Pearce L, Stechman M, Short R, Price A, Collins JT, Bruce E, Einarsson A, Rickard F, Mitchell E, Holloway M, Hesford J, Barlow-Pay F, Clini E, Myint PK, et al. 2020. The effect of frailty on survival in patients with COVID-19 (cope): a multicentre, european, observational cohort study. *The Lancet. Public Health* **5**:e444–e451. DOI: https://doi.org/10.1016/S2468-2667(20)30146-8, PMID: 32619408

Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, Lindsley RC, Mermel CH, Burtt N, Chavez A, Higgins JM, Moltchanov V, Kuo FC, Kluk MJ, Henderson B, Kinnunen L, Koistinen HA, Ladenvall C, Getz G, Correa A, et al. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *The New England Journal of Medicine* **371**:2488–2498. DOI: https://doi.org/10.1056/NEJMoa1408617, PMID: 25426837

Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E, McConkey M, Gupta N, Gabriel S, Ardissino D, Baber U, Mehran R, Fuster V, Danesh J, Frossard P, Saleheen D, Melander O, Sukhova GK, Neuberg D, Libby P, et al. 2017. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *The New England Journal of Medicine* **377**:111–121. DOI: https://doi.org/10.1056/NEJMoa1701719, PMID: 28636844

Jaiswal S, Ebert BL. 2019. Clonal hematopoiesis in human aging and disease. *Science* **366**:eaan4673. DOI: https://doi.org/10.1126/science.aan4673, PMID: 31672865

Kabat GC, Kim MY, Manson JE, Lessin L, Lin J, Wassertheil-Smoller S, Rohan TE. 2017. White blood cell count and total and cause-specific mortality in the women's health initiative. *American Journal of Epidemiology* **186**:63–72. DOI: https://doi.org/10.1093/aje/kww226, PMID: 28369251

Kar SP, Quiros PM, Gu M, Jiang T, Mitchell J, Langdon R, Iyer V, Barcena C, Vijayabaskar MS, Fabre MA, Carter P, Petrovski S, Burgess S, Vassiliou GS. 2022. Genome-wide analyses of 200,453 individuals yield new insights into the causes and consequences of clonal hematopoiesis. *Nature Genetics* **54**:1155–1166. DOI: https://doi.org/10.1038/s41588-022-01121-z, PMID: 35835912

Kristiansen M, Langerød A, Knudsen GP, Weber BL, Børresen-Dale AL, Orstavik KH. 2002. High frequency of skewed X inactivation in young breast cancer patients. *Journal of Medical Genetics* **39**:30–33. DOI: https://doi.org/10.1136/jmg.39.1.30, PMID: 11826021

Kristiansen M, Knudsen GPS, Maguire P, Margolin S, Pedersen J, Lindblom A, Ørstavik KH. 2005. High incidence of skewed X chromosome inactivation in young patients with familial non-BRCA1/BRCA2 breast cancer. *Journal of Medical Genetics* **42**:877–880. DOI: https://doi.org/10.1136/jmg.2005.032433, PMID: 15879497

Libby P, Nahrendorf M, Swirski FK. 2016. Leukocytes link local and systemic inflammation in ischemic cardiovascular disease: an expanded "cardiovascular continuum". *Coll Cardiol* **67**:1091–1103. DOI: https://doi.org/10.1016/j.jacc.2015.12.048

Ligthart S, Vaez A, Võsa U, Stathopoulou MG, de Vries PS, Prins BP, Van der Most PJ, Tanaka T, Naderi E, Rose LM, Wu Y, Karlsson R, Barbalic M, Lin H, Pool R, Zhu G, Macé A, Sidore C, Trompet S, Mangino M, et al. 2018. Genome analyses of > 200,000 individuals identify 58 loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. *The American Journal of Human Genetics* **103**:691–706. DOI: https://doi.org/10.1016/j.ajhg.2018.09.009, PMID: 30388399

López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell* **153**:1194–1217. DOI: https://doi.org/10.1016/j.cell.2013.05.039, PMID: 23746838

Lose F, Duffy DL, Kay GF, Kedda MA, Spurdle AB, Bowtell D, Chenevix-Trench G, Green A, Webb P, DeFazio A, Gertig D, Traficante N, Moore S, Hung J, Fereday S, Harrap K, Sadkowsky T, Mellon A, Robertson R, Zeps N. 2008. Skewed X chromosome inactivation and breast and ovarian cancer status: evidence for X-linked modifiers of BRCA1. *Journal of the National Cancer Institute* **100**:1519–1529. DOI: https://doi.org/10.1093/jnci/djn345, PMID: 18957670

Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Assimes TL, Ferrucci L, Horvath S. 2019. Dna methylation grimage strongly predicts lifespan and healthspan. *Aging* **11**:303–327. DOI: https://doi.org/10.18632/aging.101684, PMID: 30669119

Lyon MF. 1961. Gene action in the X-chromosome of the mouse (*Mus musculus*). *Nature* **190**:372–373. DOI: https://doi.org/10.1038/190372a0, PMID: 13764598

Madjid M, Awan I, Willerson JT, Casscells SW. 2004. Leukocyte count and coronary heart disease: implications for risk assessment. *Journal of the American College of Cardiology* **44**:1945–1956. DOI: https://doi.org/10.1016/j.jacc.2004.07.056, PMID: 15542275

Manoukian S, Verderio P, Tabano S, Colapietro P, Pizzamiglio S, Grati FR, Calvello M, Peissel B, Burn J, Pensotti V, Allemani C, Sirchia SM, Radice P, Miozzo M. 2013. X chromosome inactivation pattern in BRCA gene mutation carriers. *European Journal of Cancer* **49**:1136–1141. DOI: https://doi.org/10.1016/j.ejca.2012.10.013, PMID: 23146957

Mengel-From J, Thinggaard M, Christiansen L, Vaupel JW, Orstavik KH, Christensen K. 2012. Skewed X inactivation and survival: a 13-year follow-up study of elderly twins and singletons. *European Journal of Human Genetics* **20**:361–364. DOI: https://doi.org/10.1038/ejhg.2011.215, PMID: 22146940

Nyhan WL, Bakay B, Connor JD, Marks JF, Keele DK. 1970. Hemizygous expression of glucose-6-phosphate dehydrogenase in erythrocytes of heterozygotes for the lesch-nyhan syndrome. *PNAS* **65**:214–218. DOI: https://doi.org/10.1073/pnas.65.1.214, PMID: 5263751

Pang WW, Price EA, Sahoo D, Beerman I, Maloney WJ, Rossi DJ, Schrier SL, Weissman IL. 2011. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *PNAS* **108**:20012–20017. DOI: https://doi.org/10.1073/pnas.1116110108, PMID: 22123971

Patel KV, Ferrucci L, Ershler WB, Longo DL, Guralnik JM. 2009. Red blood cell distribution width and the risk of death in middle-aged and older adults. *Archives of Internal Medicine* **169**:515–523. DOI: https://doi.org/10.1001/archinternmed.2009.11, PMID: 19273783

Salminen A, Kaarniranta K, Kauppinen A. 2012. Inflammaging: disturbed interplay between autophagy and inflammasomes. *Aging* **4**:166–175. DOI: https://doi.org/10.18632/aging.100444, PMID: 22411934

Saraiva M, O'Garra A. 2010. The regulation of IL-10 production by immune cells. *Nature Reviews. Immunology* **10**:170–181. DOI: https://doi.org/10.1038/nri2711, PMID: 20154735

Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. 2008. A standard procedure for creating A frailty index. *BMC Geriatrics* **8**:24. DOI: https://doi.org/10.1186/1471-2318-8-24, PMID: 18826625

Skrzeczyńska-Moncznik J, Bzowska M, Loseke S, Grage-Griebenow E, Zembala M, Pryjma J. 2008. Peripheral blood cd14high CD16+ monocytes are main producers of IL-10. *Scandinavian Journal of Immunology* **67**:152–159. DOI: https://doi.org/10.1111/j.1365-3083.2007.02051.x, PMID: 18201370

Struewing JP, Pineda MA, Sherman ME, Lissowska J, Brinton LA, Peplonska B, Bardin-Mikolajczak A, Garcia-Closas M. 2006. Skewed X chromosome inactivation and early-onset breast cancer. *Journal of Medical Genetics* **43**:48–53. DOI: https://doi.org/10.1136/jmg.2005.033134, PMID: 15923273

Tam BT, Morais JA, Santosa S. 2020. Obesity and ageing: two sides of the same coin. *Obesity Reviews* **21**:e12991. DOI: https://doi.org/10.1111/obr.12991, PMID: 32020741

Therneau T. 2021. A package for survival analysis in R. R-Project. https://CRAN.R-project.org/package=survival

Tonon L, Bergamaschi G, Dellavecchia C, Rosti V, Lucotti C, Malabarba L, Novella A, Vercesi E, Frassoni F, Cazzola M. 1998. Unbalanced X-chromosome inactivation in haemopoietic cells from normal women. *British*

Journal of Haematology **102**:996–1003. DOI: https://doi.org/10.1046/j.1365-2141.1998.00867.x, PMID: 9734650

Tukiainen T, Villani A-C, Yen A, Rivas MA, Marshall JL, Satija R, Aguirre M, Gauthier L, Fleharty M, Kirby A, Cummings BB, Castel SE, Karczewski KJ, Aguet F, Byrnes A, Lappalainen T, Regev A, Ardlie KG, Hacohen N, MacArthur DG, et al. 2016. Landscape of X chromosome inactivation across human tissues. *bioRxiv*. DOI: https://doi.org/10.1101/073957

van Dijk JP, Heuver L, Stevens-Linders E, Jansen JH, Mensink E, Raymakers RAP, de Witte T. 2002. Acquired skewing of lyonization remains stable for a prolonged period in healthy blood donors. *Leukemia* **16**:362–367. DOI: https://doi.org/10.1038/sj.leu.2402379, PMID: 11896539

Verdi S, Abbasian G, Bowyer RCE, Lachance G, Yarand D, Christofidou P, Mangino M, Menni C, Bell JT, Falchi M, Small KS, Williams FMK, Hammond CJ, Hart DJ, Spector TD, Steves CJ. 2019. Twinsuk: the UK adult twin registry update. *Twin Research and Human Genetics* **22**:523–529. DOI: https://doi.org/10.1017/thg.2019.65, PMID: 31526404

Vickers MA, McLeod E, Spector TD, Wilson IJ. 2001. Assessment of mechanism of acquired skewed X inactivation by analysis of twins. *Blood* **97**:1274–1281. DOI: https://doi.org/10.1182/blood.v97.5.1274, PMID: 11222370

Wickham H. 2016. Ggplot2: elegant graphics for data analysis. Ggplot2. https://ggplot2.tidyverse.org

Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, McMichael JF, Schmidt HK, Yellapantula V, Miller CA, Ozenberger BA, Welch JS, Link DC, Walter MJ, Mardis ER, Dipersio JF, Chen F, Wilson RK, Ley TJ, Ding L. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nature Medicine* **20**:1472–1478. DOI: https://doi.org/10.1038/nm.3733, PMID: 25326804

Zito A, Davies MN, Tsai PC, Roberts S, Andres-Ejarque R, Nardone S, Bell JT, Wong CCY, Small KS. 2019. Heritability of skewed X-inactivation in female twins is tissue-specific and associated with age. *Nature Communications* **10**:5339. DOI: https://doi.org/10.1038/s41467-019-13340-w, PMID: 31767861