

## **Associations between mosaic loss of sex chromosomes and incident hospitalization for atrial fibrillation in the United Kingdom**

Jungeun Lim<sup>1</sup>, Aubrey K. Hubbard<sup>2</sup>, Batel Blechter<sup>2</sup>, Jianxin Shi<sup>2</sup>, Weiyin Zhou<sup>2,3</sup>, Erikka Loftfield<sup>2</sup>, Mitchell J. Machiela<sup>2\*</sup>, Jason Y.Y. Wong<sup>1\*</sup>

1. Epidemiology and Community Health Branch, National Heart Lung and Blood Institute, Bethesda, MD, USA.

2. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA.

3. Cancer Genomics Research Laboratory, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc, Frederick, MD, USA.

### **Corresponding author:**

Jason Wong, Epidemiology and Community Health Branch, National Heart Lung and Blood Institute, 10 Center Drive, Bethesda, MD, 20892. email: [jason.wong@nih.gov](mailto:jason.wong@nih.gov)

\*Co-supervised the study

**Running title:** Loss of sex chromosomes and heart disease

**Keywords:** Mosaic loss of sex chromosomes, incident relative risk, heart disease, prospective cohort study, atrial fibrillation

## Abstract

**Background:** Mosaic loss of chromosome Y (mLOY) in leukocytes of men reflects genomic instability from aging, smoking, and environmental exposures. A similar mosaic loss of chromosome X (mLOX) occurs among women. However, the associations between mLOY, mLOX, and risk of incident heart diseases are unclear.

**Methods:** We estimated associations between mLOY, mLOX, and risk of incident heart diseases requiring hospitalization, including atrial fibrillation, myocardial infarction, ischemic heart disease, cardiomyopathy, and heart failure. We analyzed 190,613 men and 224,853 women with genotyping data from the UK Biobank. Among these participants, we analyzed 37,037 men with mLOY and 13,978 women with mLOX detected using Mosaic Chromosomal Alterations caller. Multivariable Cox regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) of each incident heart disease in relation to mLOY in men and mLOX in women. Additionally, Mendelian randomization (MR) was conducted to estimate causal associations.

**Results:** Among men, detectable mLOY was associated with elevated risk of atrial fibrillation (HR=1.06, 95%CI:1.03-1.11). The associations were apparent in both never-smokers (HR=1.07, 95%:1.01-1.14) and ever-smokers (HR=1.05, 95%CI:1.01-1.11) as well as men > and ≤60 years of age. MR analyses supported causal associations between mLOY and atrial fibrillation (HR<sub>MR-PRESSO</sub>=1.15, 95%CI:1.13-1.18). Among post-menopausal women, we found a suggestive inverse association between detectable mLOX and atrial fibrillation risk (HR=0.90, 95%CI:0.83-0.98). However, associations with mLOY and mLOX were not found for other heart diseases.

**Conclusions:** Our findings suggest that mLOY and mLOX reflect sex-specific biological processes or exposure profiles related to incident atrial fibrillation requiring hospitalization.

### **WHAT IS KNOWN?**

- A previous population study found links between death from atrial fibrillation and heart failure, and mosaic loss of chromosome Y (mLOY) in leukocytes, which is a marker of genomic instability, environmental exposures, and aging. Additionally, mLOY has been associated cross-sectionally with prevalent cardiovascular and metabolic diseases.
- A similar but less common mosaic loss of chromosome X (mLOX) occurs among women but its role in disease pathogenesis is less characterized.
- The contributions of mLOY and mLOX to risk of hospitalization for incident atrial fibrillation, heart failure syndrome, and other heart diseases are unclear.

### **WHAT THE STUDY ADDS?**

- Among men, mLOY was associated with elevated risk of incident atrial fibrillation. Further, we found suggestive evidence linking mLOX to atrial fibrillation risk among women. Taken together, mLOY and mLOX potentially reflect sex-specific factors related to the pathogenesis of atrial fibrillation.

## Introduction

The Y chromosome contains genes that play crucial roles in biological pathways involved in male sex determination, spermatogenesis, and sexual maturation. Further, genetic variants on the Y chromosome have been linked to fundamental cellular housekeeping processes including mitosis, cell cycle regulation, DNA damage detection and repair, and apoptosis<sup>1</sup>. As such, alterations to the dosage of genes located on the Y chromosome can potentially influence the pathogenesis of chronic diseases that involve these underlying biological pathways. Mosaic loss of chromosome Y (mLOY) is a common genomic alteration that manifests in a fraction of cells as men age. Depending on the study and calling methods, an estimated 7-57% of older men, generally over 65 years of age, have detectable mLOY in circulating leukocytes<sup>1-7</sup>. Variation in mLOY has both genetic and environmental components, having associations reported with cigarette smoking, outdoor air pollution, and arsenic<sup>1,4,7,8</sup>.

In population studies, mLOY has been linked to risk of various chronic diseases<sup>9-12 13 14</sup>; however, investigations into the contribution of mLOY to the pathogenesis of cardiovascular diseases (CVDs) are still in early stages<sup>14,15</sup>. In the UK Biobank, mLOY in leukocytes was found to be cross-sectionally associated with prevalent diabetes and heart disease<sup>15</sup>.

Furthermore, recent analyses found that mLOY was associated with increased risk of death from both atrial fibrillation and heart failure syndrome<sup>16</sup>. The biological mechanisms underlying these relationships with heart diseases in humans remain unclear. However, studies have suggested links between mLOY and increased total and bioavailable testosterone<sup>17</sup>, metabolic biomarkers that we recently found to be associated with incident hospitalization for heart failure<sup>18</sup>.

The X chromosome contains essential housekeeping genes responsible for basic biological function<sup>19</sup>. Among men, mosaic loss of chromosome X (mLOX) in leukocytes are poorly tolerated and rarely observed<sup>20</sup>. Among women however, mLOX in leukocytes can change the dosage and allelic balance of genes on the maternal or paternal X chromosome. Given that one X chromosome is epigenetically inactivated in female diploid cells, the biological impact of mLOX is determined by which X-chromosome is lost<sup>21</sup>. Previous analyses found that mLOX preferentially affects the inactivated X chromosome<sup>19</sup>. Recent meta-analyses of large genome-wide association studies (GWAS) conducted in multiple biobanks have identified 49 independent genetic variants associated with mLOX<sup>21</sup>. However, little is known about its influence on the risk of heart diseases.

Despite a growing body of literature on the relationships between mosaic loss of sex chromosomes and prevalent CVD and CVD-related mortality, their associations with disease incidence are unclear. Investigating the impact of mosaic loss of sex chromosomes in leukocytes on risk of incident heart diseases is important for understanding the role of genomic instability and immune function in disease pathogenesis and could potentially have translational relevance towards risk prediction. To address these gaps in knowledge, we leveraged the extensive demographic, clinical, and GWAS data from the UK Biobank. In a population free of CVD at baseline, we evaluated the prospective associations between pre-diagnostic leukocyte mLOY among men, mLOX among women, and risk of incident atrial fibrillation, myocardial infarction, ischemic heart disease, cardiomyopathy, and heart failure syndrome requiring hospitalization.

## **Methods**

### *Study design*

The UK Biobank is a prospective cohort study and has been described in detail (<http://www.ukbiobank.ac.uk/>)<sup>22,23</sup>. Briefly, the source population included adults aged 40-69 years who lived  $\leq 40$  km of 22 study assessment centers across England, Wales, and Scotland. Nearly 9.2 million people registered in the National Health Service (NHS) were mailed invitations to participate in the study, while 503,317 (5.5%) visited the assessment centers in 2006-2010 and were enrolled<sup>22</sup>. The volunteers completed touchscreen questionnaires and physical examinations, and provided blood samples for bioassays and genomic analyses. The UK Biobank is regularly updated and our dataset from August 2022 included 502,409 subjects (project number: 28072).

### *Genome-wide association study (GWAS) data*

GWAS data was available for 488,377 UK Biobank participants as described<sup>24</sup>. Briefly, genotyping was performed by the Affymetrix Research Services Laboratory (now part of Thermo Fisher Scientific) on two similar custom arrays. Among these subjects, 49,950 were scanned on the UK BiLEVE Axiom array, while the remaining 438,427 participants were scanned using the UK Biobank Axiom array. The GWAS dataset combines results from both arrays and there were 805,426 markers in the released genotype dataset. Approximately ~3% of eligible participants who consented to genetic analyses did not have sufficient extracted DNA to conduct the genotyping. Affymetrix used a custom genotype calling pipeline and the genotype data were

quality controlled (QC). Furthermore, the dataset was phased and nearly 96 million genotypes were imputed using computationally efficient methods combined with the Haplotype Reference Consortium and UK10K haplotype resources. Information on the QC pipeline and population structure and relatedness have been described in detail <sup>24</sup>.

### *Detection of mosaic loss of sex chromosomes*

Raw genotyping array intensity data were used to calculate B allele frequencies (BAFs) and log<sub>2</sub> R ratios (LRR), which were then analyzed using MoChA v2022-01-12 (<https://github.com/freeseek/mochawdl>) to detect mosaic chromosomal alteration (mCA) regions as previously described <sup>25-27</sup>. The mCA calling was performed by Giulio Genovese (Broad Institute of MIT and Harvard). MoChA uses Viterbi hidden Markov models to integrate LRR and BAFs, leveraging haplotype information to detect subtle imbalances between maternal and paternal allelic fractions in a cell population. SHAPEIT4 software was used for phasing to infer haplotypes. LRR was used to determine the status of events (i.e., gain, loss, copy neutral loss of heterozygosity). We focused on subjects with a sample call rate  $\geq 0.97$  and `baf_auto`  $\leq 0.03$  to ensure high-quality samples for subsequent analysis.

Post-processing of the mCA data to generate the mLOX and mLOY calls was performed at the National Cancer Institute. For mLOY calling, we identified male subjects exhibiting mLOY based on the established criteria outlined at (<https://github.com/freeseek/mocha>). These criteria included a minimum event size  $> 2\text{MB}$  and a relative copy number (`rel_cov`)  $< 2.5$  in the pseudoautosomal region 1 (PAR1) region. The proportion of cells with Y-chromosome loss

(“mosaic cell fraction”, CF) was estimated using BAF values in the mLOY region. To improve mLOY detection, mosaic fraction estimates for Y-chromosome events were derived using BAF deviation on the pseudoautosomal 1 region. For the mLOX calling, we identified female subjects exhibiting mLOX based on the standard criteria specified at (<https://github.com/freeseek/mocha>), including a minimum event size >100MB and a relative copy number (rel\_cov) < 2.5 on the X chromosome. The CF with X-chromosome loss was estimated using BAF values in the mLOX region. Mosaic CF has emerged as an important consideration in mCA analyses, as increased proportions reflect greater white blood cell aberrations and stronger disease associations have been found at CFs >10%<sup>28</sup>.

#### *Ascertainment of heart disease cases*

Heart disease cases were defined using in-patient hospital diagnoses coded according to International Classification of Disease (ICD) versions 9 and 10. Atrial fibrillation was defined using ICD-10 (I48, I48.0, I48.1, I48.2, I48.3, I48.4, I48.9) and ICD-9 (427.3). Myocardial infarction was defined using ICD-10 (I21, I22, I23) and ICD-9 (412). Ischemic heart disease / angina pectoris was defined using ICD-10 (I20, I24, I25) and ICD-9 (411.1, 413.0, 413.1, 414.0, 414.8, 414.9). Cardiomyopathy was defined using ICD-10 (I25.5, I42.0, I42.5, I42.9, I43) and ICD-9 (425.4). Heart failure syndrome was defined using ICD-10 (I42.8, I42.1, I50, I50.1, I50.9, I11.0, I13.0, I13.2) and ICD-9 (428.0, 428.1).

We conducted sensitivity analysis using a smaller subgroup of UK Biobank participants (n~229,972, 45% of study population) in which preliminary raw primary care data were collected



([https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/primary\\_care\\_data.pdf](https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/primary_care_data.pdf)). These primary care data presumably capture more moderate cases that did not initially require hospitalization. Importantly, there is no national system for collecting or sharing primary care data and the UK Biobank has worked with various data suppliers and other intermediaries to obtain primary care data. Participants provided written consent for linkage of primary care data to their health-related records. Here for example, we defined atrial fibrillation using Clinical Terms Version 3 codes that correspond to a text search for “atrial fibrillation”. In the subset with primary care data, we analyzed incident outcomes as: 1) general practitioner diagnosed specific heart disease, 2) hospitalization for specific heart disease, and 3) combined general practitioner and hospitalization for specific heart disease.

#### *Inclusion / exclusion criteria*

Among the 502,409 participants at baseline, we excluded 372 participants with discrepancy between genetic and self-reported sex, 40 who dropped out of the study, 2257 with prevalent heart failure syndrome and 23,199 with prevalent CVD, which left 476,541 participants. Among these subjects, 463,082 had genotype data for detecting mLOY or mLOX (Table 1) and were considered for further analysis.

#### *Follow-up*

For each participant, the prospective follow-up started at the date of visit to the assessment center in 2006-2010 and ended at the date of first incident diagnosis (hospitalization) of heart disease,

death, or administrative censoring (i.e., September 20<sup>th</sup>, 2021, for England and Wales and October 31<sup>st</sup>, 2021, for Scotland), whichever came first. Vital status and the primary underlying cause of death for participants were provided by the NHS Information Centre and the NHS Central Register. The UK Biobank study was approved by the National Information Governance Board for Health and Social Care and the NHS North West Multicenter Research Ethics Committee. All participants provided electronic informed consent.

### *Main analyses*

Separate multivariable Cox regression models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI) of incident hospitalization for atrial fibrillation, myocardial infarction, ischemic heart disease / angina pectoris, cardiomyopathy, and heart failure syndrome in relation to mLOY (detectable vs. not detectable) among men and mLOX (detectable vs. not detectable) among women. For mLOY analyses among men, we adjusted for potential confounders including study assessment center, age at recruitment (continuous), self-identified race/ethnicity to indirectly adjust for culture-specific dietary patterns, environmental factors, social determinants of health, etc. (European, Black/African, mixed, South Asian, and other groups), 10 principal components for genetic ancestry, educational attainment (i.e., qualifications), smoking status (never, former, current), body mass index (BMI; <18.5, ≥18.5 to <25, ≥25 to <30, ≥30 to <35, and ≥35 kg/m<sup>2</sup>), material deprivation (Townsend deprivation index, socioeconomic status, continuous), alcohol intake (never, former, current occasional, current <1 drink/day, current 1-3 drinks/day, current >3 drinks/day, unknown), self-reported diabetes status (none, diabetic,

unknown), glycated hemoglobin (HbA1c, mmol/mol, continuous), hypertension status based on average systolic and diastolic blood pressure at baseline (normal, elevated, stage 1 and 2 hypertension, hypertensive crisis, and unknown), total white blood cell count at enrollment, and any cancer diagnosis at enrollment. For mLOX analyses among women, we adjusted for the same covariates as men in addition to oral contraceptive use (never and ever), hormone replacement therapy use (never and ever), and menopausal status (pre-menopause, natural post-menopause, artificial menopause, and missing/unknown). We conducted additional analyses stratified by leukocyte mosaic CFs, smoking status (never and ever), age groups ( $>$  and  $\leq 60$  years), menopausal status (pre- and post-menopause) among women, as well as a sensitivity analysis restricted to European genetic ancestry participants. We analyzed participants with complete data on independent variables. Overall and within stratified and subgroup analyses, p-values  $< 0.05$  were considered noteworthy and those below a Bonferroni-corrected  $\alpha$ -threshold of 0.01 were considered statistically significant (i.e., 0.05 divided by a family of 5 tests corresponding to each heart disease outcome).

#### *Attenuation and Sensitivity Analyses*

To tightly control potential residual confounding from smoking in the main analyses, we adjusted for a 27-category variable for smoking history and intensity<sup>12,15</sup>. Additionally, to account for non-linear relationships with age, we included polynomial terms for age in the models.

#### *Mendelian Randomization analyses*

Among the measured mLOY, mLOX, and heart disease associations found to be statistically significant in the Cox regression analyses, we further assessed potential causal associations using several established Mendelian randomization (MR) methods. First, we used the inverse variance method (IVW), an effective approach when all included genetic variants are valid instrumental variables (IVs)<sup>29</sup>. However, when horizontal pleiotropy is present, the estimate from IVW approach can be potentially biased with over-estimation<sup>29,30</sup>, and we therefore employed additional approaches as sensitivity analyses. We used the weighted median approach, which has been shown to produce valid causal estimates if the invalid IVs account for <50% of the weight of the studies<sup>29</sup>. Next, we used MR-Egger, which unlike other MR methods, estimates the dose-response relationship between the genetic associations with mLOY and those with the heart disease outcome<sup>31</sup>. A non-zero intercept from MR-Egger regression indicates potential horizontal pleiotropy, whereas the slope reflects the effect of mLOY on heart disease adjusted for horizontal pleiotropy<sup>32</sup>. Next, we used Mendelian Randomization Pleiotropy Residual Sum and Outlier (MR-PRESSO), an approach that effectively accounts for horizontal pleiotropy by identifying IVs that are outliers<sup>30</sup>. By doing so, the method prevents potential horizontal pleiotropy and corrects the causal estimates<sup>30</sup>. Lastly, we used Mendelian Randomization using a Robust Adjusted Profile Score (MR-RAPS), which maximizes the profile likelihood of the ratio estimates while accounting for weak instrument bias, pleiotropy and extreme outliers<sup>33</sup>. We used the MendelianRandomization, MR-PRESSO and mr.raps packages in R statistical software (version 4.2.2) to conduct MR analyses with summary GWAS data from a previously published study and newly conducted analyses<sup>1,34</sup> (*Online Supplementary Data File*).

## Results

### *Baseline characteristics*

The study population characteristics at baseline are shown in Table 1. Here, 19.4% of the male study participants had detectable mLOY, while 6.1% of the female study participants had detectable mLOX. On average, men with detectable mLOY were older (61.6 vs. 55.1 years), had higher proportions of former (42.6% vs. 35.8%) or current smokers (16.7% vs. 11.5%), had lower proportions of non-Europeans, and had lower proportions of college/university degree holders (30.4% vs. 35.7%) (Table 1). Similarly, among women, those with detectable mLOX were older (60.6 vs. 55.9 years) on average, had higher proportions of former smokers (34.09% vs. 30.90%), had lower proportions of non-Europeans, and had lower proportions of college/university degree holders (25.6% vs. 32.0%) (Table 1).

### *Incident heart disease cases*

Among men, there were 14,459 incident cases (hospitalizations) of atrial fibrillation, 6,007 cases of myocardial infarction, 19,507 cases of ischemic heart disease, 1,017 cases of cardiomyopathy, and 5,982 cases of heart failure syndrome during the follow-up (Supplementary Table 1). Among women, there were 9,094 incident cases of atrial fibrillation, 2,609 cases of myocardial infarction, 11,322 cases of ischemic heart disease, 565 cases of cardiomyopathy, and 3,896 cases of heart failure syndrome during the follow-up (Supplementary Table 1). The ages at first diagnosis (hospitalization) and the follow-up time to first diagnosis for each heart disease are shown in Supplementary Table 1.

### *Associations of age and smoking with mLOY and mLOX*

We conducted unadjusted analyses with age, smoking status, and sex chromosome loss at baseline to check the consistency with previous studies. Age was associated with increased odds of mLOY in men (OR=1.13 per year, 95%CI: 1.13-1.13,  $p<0.0001$ ) and mLOX in women (OR=1.09 per year, 95%CI: 1.09-1.10,  $p<0.0001$ ). As expected, among men, former-smokers (OR=1.54, 95%CI:1.50-1.58,  $p<0.0001$ ) and current-smokers (OR=1.88, 95%CI: 1.82-1.95,  $p<0.0001$ ) had increased odds of mLOY compared to never-smokers. Among post-menopausal women, former-smoking (OR=1.07, 95%CI: 1.03-1.11,  $p=0.0014$ ) was associated with increased odds of mLOX, but not current-smoking (OR=0.97, 95%CI: 0.90-1.04,  $p=0.3865$ ), which was similar to previous analyses <sup>21</sup>.

### *mLOY among men and future risk of heart disease*

Among men, having detectable mLOY in leukocytes was significantly associated with elevated risk of incident atrial fibrillation requiring hospitalization (HR=1.06, 95%CI:1.03-1.11,  $p=0.0011$ ; Table 2), even after accounting for multiple comparisons. Further, the association with atrial fibrillation strengthened non-monotonically as the mosaic CF increased above 20% (HR<sub>CF=0.20-<0.30</sub>=1.20, 95%:1.07-1.35; HR<sub>CF=0.30-<0.40</sub>=0.95, 95%:1.12-1.32; HR<sub>CF=0.40-<0.50</sub>=1.38, 95%:1.13-1.69,  $p=0.0015$ ; HR<sub>CF>0.50</sub>=1.44, 95%:1.15-1.80,  $p=0.0015$ ; Figure 1). The association between detectable mLOY and atrial fibrillation was consistent across subgroups including those of European genetic ancestry (HR=1.08, 95%CI:1.04-1.12,  $p=0.0002$ ), >60 years of age (HR=1.08, 95%CI:1.04-1.13,  $p=0.0003$ ), and ≤60 years of age (HR=1.07, 95%CI:1.02-1.11,

p=0.0021; Supplementary Table 2). The atrial fibrillation associations among never-smokers (HR=1.07, 95%CI:1.01-1.14, p=0.0294) and ever-smokers (HR=1.05, 95%CI:1.01-1.11, p=0.0298) were apparent, but non-significant after Bonferroni-correction. (Supplementary Table 2). Consistent associations were not found for myocardial infarction, ischemic heart disease / angina pectoris, cardiomyopathy, and heart failure syndrome in the overall (Table 2) and stratified analyses (Supplementary Table 2)

### *Sensitivity and attenuation analyses*

To tightly control potential residual confounding from smoking, we adjusted for a 27-category variable for smoking history and intensity<sup>12,15</sup>. Here, the mLOY-atrial fibrillation estimates were not considerably changed (HR=1.06, 95%CI:1.02-1.10, p=0.0016). Additionally, we included a quadratic term for age; however, this parameter was non-significant (p=0.19) and the main mLOY-atrial fibrillation association remained statistically significant (HR=1.06, 95%CI:1.02-1.10, p=0.0014).

When analyzing a subgroup of 87,041 participants who had primary care data and met our inclusion criteria, we found that the association between mLOY and general practitioner diagnosed incident atrial fibrillation had a similar effect size but was non-significant (605 cases/16,749 with mLOY; 1470 cases/70,292 without mLOY; HR=1.08, 95%CI: 0.96-1.20, p=0.19) compared to incident hospitalization for atrial fibrillation (1849 cases/16,749 with mLOY; 4538 cases/70,292 without mLOY; HR=1.08, 95%CI: 1.02-1.14, p=8.1x10<sup>-3</sup>). However, there was no evidence that these estimates were different (p-difference  $Z_{score\ sign}=1.00$ ). When

combining general practitioner and hospitalization outcome ascertainment, the association between mLOY and atrial fibrillation risk was statistically significant (1933 cases/16,749 with mLOY; 4746 cases/70,292 without mLOY; HR=1.09, 95%CI: 1.03-1.16,  $p=2.1 \times 10^{-3}$ ).

### *Mendelian Randomization, mLOY and atrial fibrillation risk*

We conducted MR analyses to assess potential causal associations between mLOY and atrial fibrillation risk among men. Using IVW, we found evidence supporting a causal relationship between mLOY and atrial fibrillation (HR<sub>IVW</sub>=1.15, 95% CI:1.12-1.19,  $p=1.52 \times 10^{-21}$ ; Figure 2). We conducted sensitivity analyses with other MR methods and found that the weighted median (HR<sub>WM</sub>=1.09, 95% CI:1.05-1.14,  $p=1.42 \times 10^{-24}$ ; Figure 2), MR-PRESSO (HR<sub>MR-PRESSO</sub>=1.15, 95% CI:1.13-1.18,  $p=5.22 \times 10^{-5}$ ; Figure 2) and MR-RAPS approaches (HR<sub>MR</sub>=1.10, 95% CI:1.07-1.13,  $p=1.10 \times 10^{-9}$ ; Figure 2) agreed with our main finding. Conversely, the MR-Egger slope was non-significant ( $p=0.188$ ) and the intercept was non-zero ( $\beta_0=0.011$ ,  $p<0.001$ ), suggesting detection of pleiotropy.

### *mLOX among women and future risk of heart disease*

Among women overall, having detectable mLOX in leukocytes was associated with decreased risk of atrial fibrillation requiring hospitalization before accounting for multiple testing; however, this association was marginally non-significant after Bonferroni-correction (HR=0.91, 95%CI:0.84-0.98,  $p=0.0130$ ; Table 3). The suggestive inverse association between mLOX and atrial fibrillation risk was observed only among post-menopausal women (HR=0.90,



95%CI:0.83-0.98,  $p=0.0151$ ; Table 3), who comprise most of the female participants. When further stratifying women into subgroups, the suggestive inverse association with atrial fibrillation was consistent among participants of European ancestry (HR=0.91, 95%CI:0.84-0.98,  $p=0.0145$ ), and those above and below 60 years of age at baseline (HR=0.90, 95%CI:0.83-0.98,  $p=0.0164$  and HR=0.89, 95%CI:0.82-0.96,  $p=0.0044$ , respectively) before accounting for multiple testing, but was significant only among women  $\leq 60$  years of age after Bonferroni-correction (Supplementary Table 3).

## Discussion

In a large prospective cohort study from the United Kingdom, we investigated the associations between mLOY and mLOX in pre-diagnostic leukocytes collected at baseline and risk of incident heart disease requiring hospitalization. Among men, we found that mLOY was significantly associated with elevated risk of incident atrial fibrillation. This association was greatly strengthened among participants with increased CF with mLOY and was largely consistent across subgroups defined by smoking, age, and race/ethnicity. Further, most of the MR analyses supported causal associations between mLOY and atrial fibrillation. However, consistent associations between mLOY and other heart diseases were not detected. To our knowledge, we conducted the first study of mLOX among women and risk of newly diagnosed heart disease. Here, we found suggestive evidence for inverse associations between mLOX and atrial fibrillation risk that was non-significant after accounting for multiple comparisons. When stratifying by menopausal status, the suggestive association between mLOX and atrial fibrillation was only observed among post-menopausal women. Furthermore, given that mLOX is less

frequent among younger women<sup>19</sup> and that most female UK Biobank participants were >50 years of age, we did not have sufficient sample size to make sound inferences for premenopausal women. Additionally, consistent associations were not detected between mLOX and other heart diseases among women.

Our study examines the relationships between sex chromosome mosaicism and CVD by focusing specifically on incident heart disease hospitalization, as opposed to prevalence and mortality as in other studies<sup>15,16,35</sup>. We provide population-based evidence that mLOY and mLOX could be involved (either directly or indirectly as a surrogate marker) in the biological processes leading up to atrial fibrillation occurrence. Our investigation builds upon previous studies conducted in human populations. In the UK Biobank, Loftfield et al. (2018) reported cross-sectional associations between prevalent heart disease and stroke and increased leukocyte mLOY at baseline<sup>15</sup>. Further, they reported no evidence for associations between mLOY and mortality from cardiovascular disease, ischemic heart disease, and stroke<sup>15</sup>. Additionally, a recent study by Sano et al. (2022) found evidence for associations between mLOY and death caused by diseases of the circulatory system including hypertensive heart disease, atrial fibrillation, as well as heart failure syndrome<sup>16</sup>.

We note that our outcome ascertainment using national hospital registries potentially enriches for more serious cases of heart disease, which can be non-fatal or fatal. As such, we conducted analyses using a subset of participants with preliminary primary care data, in which the cases had moderate, non-fatal incident heart disease. Here, we found that the associations between mLOY and general practitioner diagnosed atrial fibrillation were similar in magnitude to the

hospitalization data, but was not statistically significant. This was likely due to limited statistical power from reduced case numbers. Important for consideration, these preliminary primary care data have not been validated or adjudicated and we include these only for exploratory purposes and sensitivity analyses. When combining general practitioner diagnosed and hospitalization cases, the mLOY-atrial fibrillation association was statistically significant.

Our MR analyses using IVW, weighted median, MR-PRESSO, and MR-RAPS strongly supported causal associations between mLOY and atrial fibrillation and had similar causal effect estimates. MR-PRESSO is a newer method that effectively accounts for outlying genetic instruments that show evidence of horizontal pleiotropy; however, no outliers were detected in these analyses. Similarly, the newest method, MR-RAPS, accounts for weak instrument bias, pleiotropy and extreme outliers. However, we found that dose-response relationship between the genetic associations with mLOY and atrial fibrillation was non-significant using the older MR-Egger regression, and the non-zero intercept suggested some pleiotropy. We note that MR-Egger has been reported to be less statistically efficient and subject to biased or skewed estimates<sup>36</sup>, particularly when the InSIDE (INstrument Strength Independent of Direct Effect; independence between the Egger regression intercept and slope) assumption is violated<sup>31</sup>.

The precise biological roles of mLOY and mLOX in the pathogenesis of heart disease in humans is unclear, but mechanistic studies using mouse models provide some clues. For example, Sano et al. (2022) demonstrated that mLOY promotes fibrotic phenotypes ('scarring') in cardiac murine tissues which potentially impairs cardiac function such as tissue elasticity and electrical signal propagation<sup>16</sup>. Further, Horitani et al. (2024) found evidence that mLOY in mice may

disrupt the ubiquitously transcribed tetratricopeptide repeat containing, Y-linked (Uty) gene, leading to epigenetic dysfunction and differentiation of profibrotic macrophages<sup>37</sup>. Indeed, this mechanistic evidence from mice is consistent with the association we found between mLOY and atrial fibrillation risk and is consistent with the pathology of atrial fibrillation. Furthermore, evidence from sequencing data suggests that mLOY is genetically linked to clonal hematopoiesis of indeterminate potential (CHIP)<sup>38-40</sup>, a mosaic expansion of hematopoietic stem cells driven by certain somatic mutations<sup>41</sup>, which has been recently associated with risk of CVDs and heart failure syndrome<sup>42-46</sup>. Additionally, it has been hypothesized that mLOY in leukocytes potentially reflects altered immunosurveillance and inflammatory processes, which can influence cardiovascular disease pathogenesis<sup>3,47</sup>.

We note that atrial fibrillation risk was positively associated with mLOY among men as expected but was potentially inversely associated with mLOX among women. The underlying reason for this potential inverse association with mLOX is unknown. It is possible that mLOY and mLOX operate in a distinctly sex-specific manner in reflecting genomic instability from the cumulative burden of age, genetic susceptibility, and environmental exposures.

Our study had numerous strengths. First, we used MoChA to call chromosomal mosaicisms from GWAS data, which is more sensitive compared to previous calling approaches<sup>17</sup>. Second, the prospective cohort design strengthened our inferences and is robust against disease-effect bias of pre-diagnostic leukocyte mLOY and mLOX (i.e., “reverse causation”). Third, we analyzed a large sample size of middle-aged and older adults which improved statistical power to detect modest associations.

Our study had some limitations. First, even though our female sample size was large, we had limited power to detect modest associations, particularly among younger pre-menopausal women. However, we did detect suggestive associations for mLOX among post-menopausal women which could serve as a foundation for future analyses in larger pooled studies. Second, although we thoroughly adjusted for measured confounders related to chromosomal mosaicisms and heart disease, we cannot discount the possibility of unmeasured or residual confounding. However, our findings were consistent across never- and ever-smokers and age groups; supported by MR analyses; robust to adjustment for detailed smoking history; and consistent with previous studies<sup>16</sup>. Lastly, there is the possibility that healthy volunteer bias in the UK Biobank could limit generalizability; however, this does not affect the internal validity of our mechanistic study.

In conclusion, mLOY was significantly associated with elevated risk of atrial fibrillation among men. Conversely, we found suggestive evidence of an inverse association between mLOX and atrial fibrillation risk among post-menopausal women. Associations were not detected for other heart diseases. Caution is recommended when interpreting the findings, particularly for mLOX among women. Although the precise biological role of sex chromosome mosaicisms in heart disease pathogenesis remains unclear, our findings support biologic processes related to or tagged by leukocyte mLOY and mLOX are related to atrial fibrillation occurrence in human populations. If our findings are confirmed across diverse populations, they support further consideration of mLOY in atrial fibrillation risk prediction and stratification analyses among men.

## Acknowledgements

This study was supported by intramural funding from the National Cancer Institute and National Heart, Lung, and Blood Institute. We thank Drs. Qing Lan and Nathaniel Rothman for their scientific support. We extend our appreciation to Zhonghua Liu, Haoyu Zhang, and Peter Kraft for their advice on interpreting the MR analyses. Additionally, we thank Lisa Finkelstein, Jillian Varonin, and the UK Biobank Access Team for helping us with international data agreements. We declare no conflicts of interest. Lastly, we thank Giulio Genovese and John R.B. Perry for contributing to the calling of chromosomal mosaicisms.

## References

1. Thompson DJ, Genovese G, Halvardson J, Ulirsch JC, Wright DJ, Terao C, Davidsson OB, Day FR, Sulem P, Jiang Y, et al. Genetic predisposition to mosaic Y chromosome loss in blood. *Nature*. 2019;575:652-657. doi: 10.1038/s41586-019-1765-3
2. Garcia-Gonzalez P, de Rojas I, Moreno-Grau S, Montreal L, Puerta R, Alarcon-Martin E, Quintela I, Orellana A, Andrade V, Adami PVM, et al. Mendelian Randomisation Confirms the Role of Y-Chromosome Loss in Alzheimer's Disease Aetiopathogenesis in Men. *Int J Mol Sci*. 2023;24. doi: 10.3390/ijms24020898
3. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, Sandgren J, Diaz de Stahl T, Zaghlool A, Giedraitis V, et al. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nat Genet*. 2014;46:624-628. doi: 10.1038/ng.2966

4. Wong JYY, Margolis HG, Machiela M, Zhou W, Odden MC, Psaty BM, Robbins J, Jones RR, Rotter JI, Chanock SJ, et al. Outdoor air pollution and mosaic loss of chromosome Y in older men from the Cardiovascular Health Study. *Environ Int*. 2018;116:239-247. doi: 10.1016/j.envint.2018.04.030
5. Zhou W, Machiela MJ, Freedman ND, Rothman N, Malats N, Dagnall C, Caporaso N, Teras LT, Gaudet MM, Gapstur SM, et al. Mosaic loss of chromosome Y is associated with common variation near TCL1A. *Nat Genet*. 2016;48:563-568. doi: 10.1038/ng.3545
6. Forsberg LA, Halvardson J, Rychlicka-Buniowska E, Danielsson M, Moghadam BT, Mattisson J, Rasi C, Davies H, Lind L, Giedraitis V, et al. Mosaic loss of chromosome Y in leukocytes matters. *Nat Genet*. 2019;51:4-7. doi: 10.1038/s41588-018-0267-9
7. Demanelis K, Delgado DA, Tong L, Jasmine F, Ahmed A, Islam T, Parvez F, Kibriya MG, Graziano JH, Ahsan H, et al. Somatic loss of the Y chromosome is associated with arsenic exposure among Bangladeshi men. *Int J Epidemiol*. 2023;52:1035-1046. doi: 10.1093/ije/dyac176
8. Dumanski JP, Rasi C, Lonn M, Davies H, Ingelsson M, Giedraitis V, Lannfelt L, Magnusson PK, Lindgren CM, Morris AP, et al. Mutagenesis. Smoking is associated with mosaic loss of chromosome Y. *Science*. 2015;347:81-83. doi: 10.1126/science.1262092
9. Brown DW, Lan Q, Rothman N, Pluta J, Almstrup K, Dalgaard MD, Greene MH, Grotmol T, Loveday C, Schwartz SM, et al. Genetically Inferred Telomere Length and Testicular Germ Cell Tumor Risk. *Cancer Epidemiol Biomarkers Prev*. 2021;30:1275-1278. doi: 10.1158/1055-9965.EPI-20-1775

10. Machiela MJ, Dagnall CL, Pathak A, Loud JT, Chanock SJ, Greene MH, McGlynn KA, Stewart DR. Mosaic chromosome Y loss and testicular germ cell tumor risk. *J Hum Genet.* 2017;62:637-640. doi: 10.1038/jhg.2017.20
11. Kobayashi T, Hachiya T, Ikehata Y, Horie S. Genetic association of mosaic loss of chromosome Y with prostate cancer in men of European and East Asian ancestries: a Mendelian randomization study. *Front Aging.* 2023;4:1176451. doi: 10.3389/fragi.2023.1176451
12. Loftfield E, Zhou W, Yeager M, Chanock SJ, Freedman ND, Machiela MJ. Mosaic Y Loss Is Moderately Associated with Solid Tumor Risk. *Cancer Res.* 2019;79:461-466. doi: 10.1158/0008-5472.CAN-18-2566
13. Dumanski JP, Lambert JC, Rasi C, Giedraitis V, Davies H, Grenier-Boley B, Lindgren CM, Campion D, Dufouil C, European Alzheimer's Disease Initiative I, et al. Mosaic Loss of Chromosome Y in Blood Is Associated with Alzheimer Disease. *Am J Hum Genet.* 2016;98:1208-1219. doi: 10.1016/j.ajhg.2016.05.014
14. Lin SH, Brown DW, Rose B, Day F, Lee OW, Khan SM, Hislop J, Chanock SJ, Perry JRB, Machiela MJ. Incident disease associations with mosaic chromosomal alterations on autosomes, X and Y chromosomes: insights from a phenome-wide association study in the UK Biobank. *Cell Biosci.* 2021;11:143. doi: 10.1186/s13578-021-00651-z
15. Loftfield E, Zhou W, Graubard BI, Yeager M, Chanock SJ, Freedman ND, Machiela MJ. Predictors of mosaic chromosome Y loss and associations with mortality in the UK Biobank. *Sci Rep.* 2018;8:12316. doi: 10.1038/s41598-018-30759-1
16. Sano S, Horitani K, Ogawa H, Halvardson J, Chavkin NW, Wang Y, Sano M, Mattisson J, Hata A, Danielsson M, et al. Hematopoietic loss of Y chromosome leads to cardiac



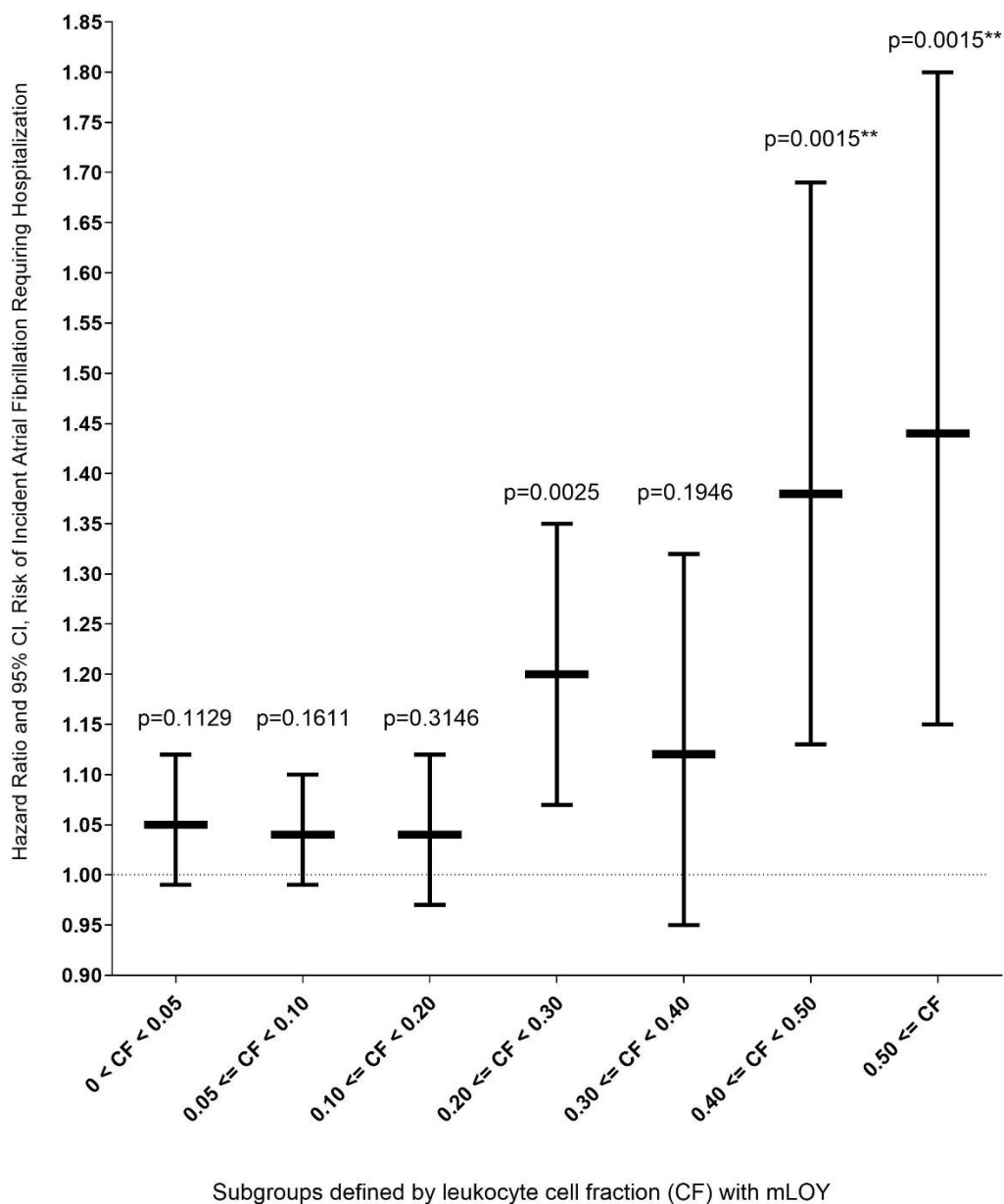
- fibrosis and heart failure mortality. *Science*. 2022;377:292-297. doi: 10.1126/science.abn3100
17. Hubbard AK, Brown DW, Zhou W, Lin SH, Genovese G, Chanock SJ, Machiela MJ. Serum biomarkers are altered in UK Biobank participants with mosaic chromosomal alterations. *Hum Mol Genet*. 2023;32:3146-3152. doi: 10.1093/hmg/ddad133
  18. Lim J, Hashemian M, Blechter B, Roger VL, Wong JYY. Pre-diagnostic free androgen and estradiol levels influence heart failure risk in both women and men: A prospective cohort study in the UK Biobank. *Eur J Heart Fail*. 2024. doi: 10.1002/ejhf.3189
  19. Machiela MJ, Zhou W, Karlins E, Sampson JN, Freedman ND, Yang Q, Hicks B, Dagnall C, Hautman C, Jacobs KB, et al. Female chromosome X mosaicism is age-related and preferentially affects the inactivated X chromosome. *Nat Commun*. 2016;7:11843. doi: 10.1038/ncomms11843
  20. Zhou W, Lin SH, Khan SM, Yeager M, Chanock SJ, Machiela MJ. Detectable chromosome X mosaicism in males is rarely tolerated in peripheral leukocytes. *Sci Rep*. 2021;11:1193. doi: 10.1038/s41598-020-80948-0
  21. Liu A, Genovese G, Zhao Y, Pirinen M, Zekavat MM, Kentistou K, Yang Z, Yu K, Vlasschaert C, Liu X, et al. Population analyses of mosaic X chromosome loss identify genetic drivers and widespread signatures of cellular selection. *medRxiv*. 2023. doi: 10.1101/2023.01.28.23285140
  22. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol*. 2017;186:1026-1034. doi: 10.1093/aje/kwx246

23. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12:e1001779. doi: 10.1371/journal.pmed.1001779
24. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562:203-209. doi: 10.1038/s41586-018-0579-z
25. Loh PR, Genovese G, Handsaker RE, Finucane HK, Reshef YA, Palamara PF, Birmann BM, Talkowski ME, Bakhoun SF, McCarroll SA, et al. Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature.* 2018;559:350-355. doi: 10.1038/s41586-018-0321-x
26. Loh PR, Genovese G, McCarroll SA. Monogenic and polygenic inheritance become instruments for clonal selection. *Nature.* 2020;584:136-141. doi: 10.1038/s41586-020-2430-6
27. Lin SH, Khan SM, Zhou W, Brown DW, Vergara C, Wolinsky SM, Martinez-Maza O, Margolick JB, Martinson JJ, Hussain SK, et al. Mosaic chromosomal alterations detected in men living with HIV and the relationship to non-Hodgkin lymphoma. *AIDS.* 2023. doi: 10.1097/QAD.0000000000003545
28. Zekavat SM, Lin SH, Bick AG, Liu A, Paruchuri K, Wang C, Uddin MM, Ye Y, Yu Z, Liu X, et al. Hematopoietic mosaic chromosomal alterations increase the risk for diverse types of infection. *Nat Med.* 2021;27:1012-1024. doi: 10.1038/s41591-021-01371-0

29. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40:304-314. doi: 10.1002/gepi.21965
30. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50:693-698. doi: 10.1038/s41588-018-0099-7
31. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32:377-389. doi: 10.1007/s10654-017-0255-x
32. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512-525. doi: 10.1093/ije/dyv080
33. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *The Annals of Statistics*. 2020;48:1742-1769, 1728.
34. Burgess SY, O. [https://cran.r-project.org/web/packages/MendelianRandomization/vignettes/Vignette\\_MR.pdf](https://cran.r-project.org/web/packages/MendelianRandomization/vignettes/Vignette_MR.pdf).
35. Haitjema S, Kofink D, van Setten J, van der Laan SW, Schoneveld AH, Eales J, Tomaszewski M, de Jager SCA, Pasterkamp G, Asselbergs FW, et al. Loss of Y Chromosome in Blood Is Associated With Major Cardiovascular Events During Follow-Up in Men After Carotid Endarterectomy. *Circ Cardiovasc Genet*. 2017;10:e001544. doi: 10.1161/CIRCGENETICS.116.001544

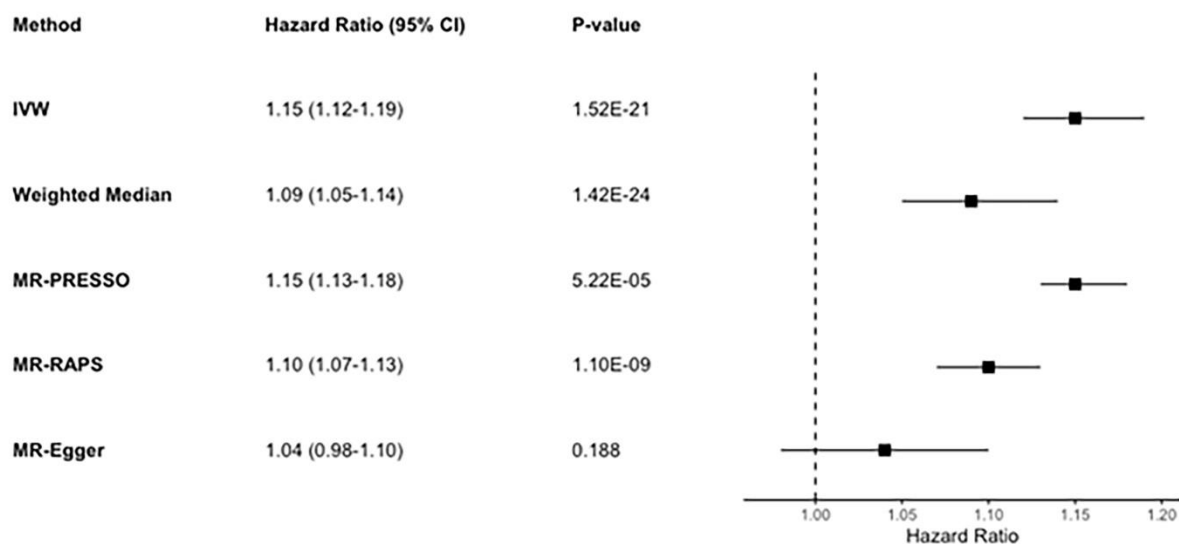
36. Qi G, Chatterjee N. Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. *Nat Commun.* 2019;10:1941. doi: 10.1038/s41467-019-09432-2
37. Horitani K, Chavkin NW, Arai Y, Wang Y, Ogawa H, Yura Y, Evans MA, Cochran JD, Thel MC, Polizio AH, et al. Disruption of the Uty epigenetic regulator locus in hematopoietic cells phenocopies the profibrotic attributes of Y chromosome loss in heart failure. *Nature Cardiovascular Research.* 2024;3:343-355. doi: 10.1038/s44161-024-00441-z
38. Jakubek YA, Reiner AP, Honigberg MC. Risk factors for clonal hematopoiesis of indeterminate potential and mosaic chromosomal alterations. *Transl Res.* 2023;255:171-180. doi: 10.1016/j.trsl.2022.11.009
39. Walsh K, Raghavachari N, Kerr C, Bick AG, Cummings SR, Druley T, Dunbar CE, Genovese G, Goodell MA, Jaiswal S, et al. Clonal Hematopoiesis Analyses in Clinical, Epidemiologic, and Genetic Aging Studies to Unravel Underlying Mechanisms of Age-Related Dysfunction in Humans. *Front Aging.* 2022;3:841796. doi: 10.3389/fragi.2022.841796
40. Ljungstrom V, Mattisson J, Halvardson J, Pandzic T, Davies H, Rychlicka-Buniowska E, Danielsson M, Lacaze P, Cavelier L, Dumanski JP, et al. Loss of Y and clonal hematopoiesis in blood-two sides of the same coin? *Leukemia.* 2022;36:889-891. doi: 10.1038/s41375-021-01456-2
41. Heimlich JB, Bick AG. Somatic Mutations in Cardiovascular Disease. *Circ Res.* 2022;130:149-161. doi: 10.1161/CIRCRESAHA.121.319809

42. Yu B, Roberts MB, Raffield LM, Zekavat SM, Nguyen NQH, Biggs ML, Brown MR, Griffin G, Desai P, Correa A, et al. Supplemental Association of Clonal Hematopoiesis With Incident Heart Failure. *J Am Coll Cardiol*. 2021;78:42-52. doi: 10.1016/j.jacc.2021.04.085
43. Reiner AP, Roberts MB, Honigberg MC, Kooperberg C, Desai P, Bick AG, Natarajan P, Manson JE, Whitsel EA, Eaton CB. Association of Clonal Hematopoiesis of Indeterminate Potential with Incident Heart Failure with Preserved Ejection Fraction. *medRxiv*. 2023. doi: 10.1101/2023.06.07.23291038
44. Jaiswal S, Libby P. Clonal haematopoiesis: connecting ageing and inflammation in cardiovascular disease. *Nat Rev Cardiol*. 2020;17:137-144. doi: 10.1038/s41569-019-0247-5
45. Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E, McConkey M, Gupta N, Gabriel S, Ardissino D, et al. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med*. 2017;377:111-121. doi: 10.1056/NEJMoa1701719
46. Libby P, Ebert BL. CHIP (Clonal Hematopoiesis of Indeterminate Potential): Potent and Newly Recognized Contributor to Cardiovascular Risk. *Circulation*. 2018;138:666-668. doi: 10.1161/CIRCULATIONAHA.118.034392
47. Swirski FK, Nahrendorf M. Cardioimmunology: the immune system in cardiac homeostasis and disease. *Nat Rev Immunol*. 2018;18:733-744. doi: 10.1038/s41577-018-0065-8



**Figure 1: Mosaic loss of chromosome Y (mLOY) in pre-diagnostic leukocytes and risk of newly diagnosed atrial fibrillation, stratified by cell fractions (CF) with mLOY.** We analyzed participants with complete genotyping and covariate data. Multivariable Cox regression models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI) of newly diagnosed (incident) atrial fibrillation in relation to mLOY (detectable vs. not detectable) among men. We adjusted for potential confounders including study assessment center, age at recruitment (continuous), self-identified race/ethnicity (European, Black/African, mixed, South Asian, and other groups), 10 principal components for genetic ancestry, educational attainment (i.e., qualifications), smoking status (never, former, current), body mass index (BMI; <18.5, ≥18.5 to <25, ≥25 to <30, ≥30 to <35, and ≥35 kg/m<sup>2</sup>), material deprivation (Townsend deprivation index, socioeconomic status, continuous), alcohol intake (never, former, current occasional, current <1

drink/day, current 1-3 drinks/day, current >3 drinks/day, unknown), diabetes status (none, diabetic, unknown), glycosylated hemoglobin (HbA1c, mmol/mol, continuous), and hypertension status based on average systolic and diastolic blood pressure at baseline (normal, elevated, stage 1 and 2 hypertension, hypertensive crisis, and unknown), total white blood cell count at enrollment, and any cancer diagnosis at enrollment. Analytic strata sizes with complete data: No detectable mLOY (comparison group): 10,245 cases/153,367 men;  $0 < CF < 0.05$ : 1090 cases/11,531 men;  $0.05 \leq CF < 0.10$ : 1649 cases/14,905 men;  $0.10 \leq CF < 0.20$ : 843 cases/6703 men;  $0.20 \leq CF < 0.30$ : 299 cases/1972 men;  $0.30 \leq CF < 0.40$ : 141 cases/939 men;  $0.40 \leq CF < 0.50$ : 98 cases/533 men;  $CF < 0.50$ : 78 cases/418 men.



**Figure 2: Mendelian Randomization analyses to assess the potential causal associations between pre-diagnostic leukocyte mLOY and risk of newly diagnosed atrial fibrillation in the UK Biobank.** CI, confidence interval; IVW, Inverse Variance Weighted; MR-PRESSO, Mendelian Randomization Pleiotropy Residual Sum and Outlier; MR-RAPS, Mendelian Randomization-Robust Adjusted Profile Score.



**Table 1: Baseline characteristics of UK Biobank study population with and without mosaic loss of sex chromosomes (n=463,082)**

Characteristic	MEN					WOMEN				
	Detectable mLOY (n=40,043) n	Mean (SD) or %	No mLOY (n=165,968) n	Mean (SD) or %	p-value	Detectable mLOX (n=15,680) n	Mean (SD) or %	No mLOX (n=241,391) n	Mean (SD) or %	p-value
Age at recruitment, years, n, mean, SD	40,043	61.6 (6.0)	165,968	55.1 (8.2)	<0.0001	15,680	60.6 (6.6)	241,391	55.9 (8.0)	
Body mass index, kg/m <sup>2</sup> , n, %										
<18.5	90	0.2	365	0.2	<0.0001	99	0.63	1,673	0.7	<0.0001
≥18.5 to <25	10,213	25.5	40,731	24.5		5,521	35.21	93,032	38.5	
≥25 to <30	20,139	50.3	81,156	48.9		6,166	39.32	87,774	36.4	
≥30 to <35	7,254	18.1	31,954	19.3		2,532	16.15	36,868	15.3	
≥35	1,659	4.1	9,257	5.6		1,127	7.19	18,237	7.6	
Missing/unknown	688	1.7	2,505	1.5		235	1.50	3,807	1.6	
Smoking status, n, %										
Never	16,133	40.3	86,637	52.2	<0.0001	8,999	57.39	144,230	59.7	<0.0001
Former	17,039	42.6	59,466	35.8		5,345	34.09	74,569	30.9	
Current	6,674	16.7	19,018	11.5		1,257	8.02	21,433	8.9	
Unknown	197	0.5	847	0.5		79	0.50	1,159	0.5	
Self-reported Race/Ethnicity/Ancestry, n, %										
European	39,023	97.5	154,824	93.3	<0.0001	15,369	98.02	227,217	94.1	<0.0001
Mixed	96	0.2	913	0.6		51	0.33	1,687	0.7	
South Asian	264	0.7	3,426	2.1		60	0.38	3,349	1.4	
Black/African	223	0.6	2,949	1.8		88	0.56	4,132	1.7	
Other groups	239	0.6	2,919	1.8		61	0.39	4,039	1.7	
Missing/unknown	198	0.5	937	0.6		51	0.33	967	0.4	
Townsend deprivation index (socioeconomic status), n,	40,004	-1.5 (3.0)	165,747	-1.2 (3.2)	<0.0001	15,666	-1.6 (2.9)	241,100	-1.4 (3.0)	

mean, SD

Education (qualifications), n,  
%

College or University degree	12,180	30.4	59,183	35.7	<0.0001	4,008	25.6	77,160	32.0	<0.0001
A levels/AS levels or equivalent	3,764	9.4	17,693	10.7		1,574	10.0	29,013	12.0	
O levels/GCSEs or equivalent	7,153	17.9	31,265	18.8		3,825	24.4	56,272	23.3	
CSEs or equivalent	1,343	3.4	10,095	6.1		672	4.3	13,243	5.5	
NVQ or HND or HNC or equivalent	4,010	10.0	14,421	8.7		701	4.5	10,702	4.4	
Other professional qualifications	2,127	5.3	6,886	4.1		1,012	6.5	13,695	5.7	
Unknown/Missing/No Answer	9,466	23.6	26,425	15.9		3,888	24.8	41,306	17.1	

Alcohol intake, n, %

Never	827	2.1	4,719	2.8	<0.0001	846	5.4	13,701	5.7	0.02
Former	1,371	3.4	5,425	3.3		551	3.5	8,471	3.5	
Current occasional	6,007	15.0	26,914	16.2		4,292	27.4	67,235	27.9	
Current <1 drink/day	7,983	19.9	33,641	20.3		4,517	28.8	70,209	29.1	
Current 1-3 drinks/day	17,629	44.0	71,479	43.1		4,891	31.2	72,539	30.1	
Current >3 drinks/day	6,150	15.4	23,251	14.0		556	3.5	8,599	3.6	
Unknown	76	0.2	539	0.3		27	0.2	637	0.3	

Self-reported diabetes status,  
n, %

None	37,368	93.3	155,269	93.6		15,036	95.9	232,269	96.2	
Diabetic	2,507	6.3	9,837	5.9	0.0018	593	3.8	8,265	3.4	
Unknown	168	0.4	862	0.5		51	0.3	857	0.4	

Glycated hemoglobin,  
HbA1c, mmol/mol, n, mean,  
SD

	38,021	36.8 (7.1)	157,700	36.1 (7.3)	<0.0001	14,889	36.3 (5.6)	229,328	35.6 (5.8)	0.05
--	--------	------------	---------	------------	---------	--------	------------	---------	------------	------

Hypertension status, n, %

Normal	2,467	6.2	13,246	8.0	<.0001	2,052	13.1	46,384	19.2	<0.0001
Elevated	3,680	9.2	17,679	10.7		1,768	11.3	28,975	12.0	
Stage 1 or 2	29,507	73.7	117,793	71.0		10,082	64.3	140,702	58.3	
Hypertensive crisis	1,201	3.0	3,398	2.0		427	2.7	4,340	1.8	

Unknown	3,188	8.0	13,852	8.3	1,351	8.6	20,990	8.7	
Menopausal Status, n, %	N/A		N/A						0.0
Pre-menopause					1,455	9.3	60,187	24.9	<0.0001
Post-menopause					11,726	74.8	143,112	59.3	
Hysterectomy					2,069	13.2	26,680	11.1	
Bilateral Oophorectomy					11	0.1	166	0.1	
Unknown/Missing									
Ever Oral Contraceptive Use, n, %	N/A		N/A						
Yes					3,691	23.5	43,834	18.2	<0.0001
No					11,941	76.2	196,495	81.4	
Ever Hormone Replacement Therapy, n, %	N/A		N/A						
Yes					7,801	49.8	151,542	62.8	<0.0001
No					7,831	49.9	88,628	36.7	

Differences between having and not having chromosomal mosaicism were compared using Chi-square tests for categorical variables and Wilcoxon signed rank tests for continuous variables. Abbreviations: Not applicable (N/A), mosaic loss of chromosome Y (mLOY), mosaic loss of chromosome X (mLOX), standard deviation (SD), hemoglobin A1C (HbA1c), advanced (A), advanced subsidiary (AS), ordinary (O), general certificate of secondary education (GCSE), certificate of secondary education (CSE), national vocational qualification (NVQ), higher national diploma (HND), and higher national certificate (HNC)

**Table 2: Mosaic loss of chromosome Y (mLOY) in pre-diagnostic leukocytes and risk of incident heart diseases in the UK Biobank**

Outcome	<i>Detectable mLOY</i>	<i>No mLOY</i>	HR	95% CI		P-value
	<i>(total n=37,037)</i>	<i>(total n=153,576)</i>		Lower	Upper	
	Cases, n	Cases, n				
Atrial fibrillation	4,203	10,256	1.06	1.03	1.11	0.0011 **
Myocardial Infarction	1,512	4,495	1.05	0.98	1.11	0.1448
Ischemic heart disease / angina pectoris	4,938	14,569	1.00	0.96	1.03	0.8706
Cardiomyopathy	260	757	1.05	0.90	1.21	0.5579
Heart failure syndrome	1,723	4,259	1.04	0.98	1.10	0.2115

\*\* P-values below a Bonferroni-corrected p-value of 0.01 (0.05 divided by a family of 5 tests). We analyzed participants with complete mLOY and covariate data.

Separate multivariable Cox regression models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI) of newly diagnosed (incident) atrial fibrillation, myocardial infarction, ischemic heart disease / angina pectoris, cardiomyopathy, and heart failure syndrome in relation to mLOY (detectable vs. not detectable) among men. We adjusted for potential confounders including study assessment center, age at recruitment (continuous), self-identified race/ethnicity (European, Black/African, mixed, South Asian, and other groups), 10 principal components for genetic ancestry, educational attainment (i.e., qualifications), smoking status (never, former, current), body mass index (BMI; <18.5, ≥18.5 to <25, ≥25 to <30, ≥30 to <35, and ≥35 kg/m<sup>2</sup>), material deprivation (Townsend deprivation index, socioeconomic status, continuous), alcohol intake (never, former, current occasional, current <1 drink/day, current 1-3 drinks/day, current >3 drinks/day, unknown), diabetes status (none, diabetic, unknown), glycated hemoglobin (HbA1c, mmol/mol, continuous), and hypertension status based on average systolic and diastolic blood pressure at baseline (normal, elevated, stage 1 and 2 hypertension, hypertensive crisis, and unknown), total white blood cell count at enrollment, and any cancer diagnosis at enrollment.

**Table 3: Mosaic loss of chromosome X (mLOX) in pre-diagnostic leukocytes from women and risk of incident heart diseases in the UK Biobank**

<b>Heart disease</b>	<b>Detectable mLOX Cases, n</b>	<b>No mLOX Cases, n</b>	<b>HR</b>	<b>95%CI Lower</b>	<b>95% CI Upper</b>	<b>P-value</b>	
<i>Overall Women (total n of subjects) †</i>	<i>13,978</i>	<i>210,875</i>					
Atrial fibrillation	748	8,346	0.91	0.84	0.98	0.0130	*
Myocardial Infarction	209	2,400	0.97	0.84	1.12	0.7012	
Ischemic heart disease / angina pectoris	927	10,395	1.00	0.93	1.07	0.9477	
Cardiomyopathy	35	530	0.79	0.56	1.11	0.1752	
Heart failure syndrome	342	3,554	0.99	0.89	1.11	0.8884	
<i>Post-menopause (total n of subjects)</i>	<i>10,743</i>	<i>131,193</i>					
Atrial fibrillation	601	6,534	0.90	0.83	0.98	0.0151	*
Myocardial Infarction	170	1,818	0.98	0.84	1.15	0.7997	
Ischemic heart disease / angina pectoris	727	7,615	1.00	0.93	1.08	0.9601	
Cardiomyopathy	28	374	0.84	0.57	1.23	0.3635	
Heart failure syndrome	272	2,750	0.98	0.86	1.11	0.7473	
<i>Pre-menopause (total n of subjects)</i>	<i>1,330</i>	<i>55,054</i>					
Atrial fibrillation	18	468	1.09	0.68	1.76	0.7160	
Myocardial Infarction	7	216	0.98	0.45	2.10	0.9527	
Ischemic heart disease / angina pectoris	40	918	1.29	0.93	1.78	0.1248	
Cardiomyopathy	2	62	1.12	0.27	4.66	0.8727	
Heart failure syndrome	9	218	1.05	0.53	2.11	0.8884	

† Includes women in pre-, post-, and artificial menopause. \*P<0.05. \*\* P-values below a Bonferroni-corrected p-value of 0.01 (0.05 divided by a family of 5 tests). We analyzed participants with complete mLOX and covariate data.

Separate multivariable Cox regression models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI) of newly diagnosed (incident) atrial fibrillation, myocardial infarction, ischemic heart disease / angina pectoris, cardiomyopathy, and heart failure syndrome in relation to mLOX (detectable vs. not detectable) among women. We adjusted for potential confounders including study assessment center, age at recruitment (continuous), self-identified race/ethnicity (European, Black/African, mixed, South Asian, and other groups), 10 principal components for genetic ancestry, educational attainment (i.e., qualifications), smoking status (never, former, current), body mass index (BMI; <18.5, ≥18.5 to <25, ≥25 to <30, ≥30 to <35, and ≥35 kg/m<sup>2</sup>), material deprivation (Townsend deprivation index, socioeconomic status, continuous), alcohol intake (never, former, current occasional, current <1 drink/day, current 1-3 drinks/day, current >3 drinks/day, unknown), diabetes status (none, diabetic, unknown), glycated hemoglobin (HbA1c, mmol/mol, continuous), and hypertension status based on average systolic and diastolic blood pressure at baseline (normal, elevated, stage 1 and 2 hypertension, hypertensive crisis, and unknown), total white blood cell count at enrollment, ever use of oral contraceptive or hormone replacement therapy, and any cancer diagnosis at enrollment.



Pre-diagnostic blood  
collected at enrollment

**n=463,082**

Mosaic loss of  
chromosome



**WOMEN**  
Decreased risk of  
atrial fibrillation  
requiring  
hospitalization



Mosaic loss of  
chromosome



**MEN**  
Increased risk of  
atrial fibrillation  
requiring  
hospitalization

