

Study of the associations between short telomeres, sex hormones and pulmonary fibrosis

Anna Duckworth PhD ^{1,2,3}, Katherine S. Ruth PhD ¹, Julia K. Prague PhD ^{3,4,5}, Anne-Marie Russell PhD ^{1,2,3}, Howard Almond ³, John Conway ³, Robin N Beaumont PhD¹, Andrew R Wood PhD¹, Susan Martin PhD¹, Professor Katie Lunnon PhD ¹, Professor Mark A. Lindsay PhD ⁶, Professor Anna Murray PhD ¹, Michael A. Gibbons PhD ^{1, 2, 3}, Jess Tyrrell PhD¹, Chris J. Scotton PhD ^{1, 2, 3}

¹ Department of Clinical & Biomedical Sciences, Faculty of Health & Life Sciences, University of Exeter, Exeter, UK

² Respiratory Medicine Department, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

³ Exeter Patients in Collaboration (EPIC) for PF, Exeter, UK

⁴ Macleod Diabetes and Endocrinology Centre, Royal Devon and Exeter National Health Service Foundation Trust, Exeter, EX2 5DW, UK.

⁵ Exeter Centre of Excellence for Diabetes Research, University of Exeter, Exeter, EX2 5DW, UK

⁶ Department of Pharmacy and Pharmacology, University of Bath, Bath, UK

Name and address for correspondence:

Dr Anna Duckworth

ad653@exeter.ac.uk

Department of Clinical & Biomedical Sciences

Faculty of Health & Life Sciences

University of Exeter

St Luke's Campus

Magdalen Road

Exeter

EX1 2LU

Tel: +44(0)1392722934

Keywords: pulmonary fibrosis, sex hormones, telomeres, testosterone, oestrogen, menopause

Abstract

Background

Pulmonary fibrosis (PF) is an incurable fibrotic lung disease with limited treatment options and a high mortality. Evidence is growing that short telomeres cause both heritable and idiopathic pulmonary fibrosis (IPF). Based on survival data, we hypothesised that sex hormones are protective against premature telomere attrition and could influence PF disease onset and/or progression.

Methods

Associations between IPF, sex hormone concentrations and measured leukocyte telomere length (LTL) were examined for unrelated UK Biobank participants of European ancestry with a diagnosis of IPF (415 females, 718 males) against controls (204,321 females, 174,254 males). Polygenic risk scores were used to explore causality between sex hormone indices, LTL and disease.

Findings

Strong associations were found between IPF and LTL. For females, higher odds of having IPF was associated with early menopause and premature ovarian failure. Menopause age correlated positively with both age of IPF diagnosis and age of death. For males, IPF prevalence and stages of disease were associated with serum bioavailable testosterone concentrations. For both sexes, evidence of lower concentrations of sex hormones was associated with shorter LTL. Genetic analysis also inferred bi-directional causal links between sex hormone binding globulin concentration, which impacts free testosterone concentration, and LTL in males.

Interpretation

Our findings suggest that higher sex hormone concentrations protect against IPF onset and progression, possibly by slowing telomere shortening. Hormonal supplementation may delay or prevent disease onset for those with telomere-associated PF risk and improve disease prognosis. This warrants further exploration in a randomised controlled trial.

Funding

Medical Research Council.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a complex and incurable fibrotic lung disease with increasing prevalence and mortality in the UK. Currently IPF accounts for around 5300 UK deaths each year¹ (62% male²) and occurs later in life when sex hormone concentrations are lower.

Circulating sex steroid concentrations are regulated by the hypothalamic-pituitary-gonadal axis in a system of negative feedback to maintain equilibrium throughout life. In men, 1.2-2.9% of total circulating testosterone (depending on age) is freely available to activate testosterone-responsive cells, 33-54% is bound with low affinity to serum albumin (which together constitute 'bioavailable' testosterone³) and the remainder is tightly bound to sex hormone binding globulin (SHBG)⁴. Similar control of oestrogen bioavailability is seen in women, where the majority is bound to albumin and SHBG. SHBG concentrations vary between, and within, individuals.

Since 2007, evidence has grown detailing the association of pulmonary fibrosis (PF) with short telomeres, the DNA sequences that protect the ends of chromosomes from aberrant repair and inter-chromosomal fusion^{5,6}. Familial clustering of PF in patients with dyskeratosis congenita (DC), a rare heritable bone marrow failure syndrome, first implicated mutations in the telomere maintenance genes including *TERT* (telomerase reverse transcriptase) and *TERC* (telomerase RNA component)⁵. Telomerase, which catalyses telomere elongation⁷, is expressed in stem cells, including lung alveolar epithelial cells, where its insufficiency precipitates a cascade of replicative senescence, aberrant wound healing and tissue fibrosis. Recent evidence suggests that short telomeres also contribute to idiopathic PF risk⁸ and associate with worse survival⁹. Using Mendelian randomisation (MR), we recently inferred that telomere length is causally linked with IPF (OR=5.81, 95%CI:3.56-9.50, $P=2.19 \times 10^{-12}$)¹⁰.

DC and congenital aplastic anaemia (also frequently associated with *TERT* and *TERC* mutations) represent severe telomere biology disorders (TBDs) and both have been treated effectively with androgens since the 1960s^{11,12,13}, although case reports are mixed¹⁴ and concern exists regarding increased risk of malignancy due to the association between upregulated telomerase and tumourigenesis¹⁵. However, a recent review of fibrosis in the kidney, liver, lung and heart concluded that sex-related differences in the pathogenesis strongly suggest a role for sex hormones in modulating disease progression¹⁶. Evidence from androgen deprivation therapy in prostate cancer suggests that the detrimental impact of reduced sex hormone levels is restricted to those at risk for TBDs¹⁷.

In general, women have longer telomeres¹⁸, lower PF prevalence¹⁹ and slower PF progression than men²⁰. A study of 101 Chinese male IPF patients versus 53 controls matched for age and smoking history found that patients had significantly lower free testosterone (freeT; 210 ± 8 pmol/l versus 281 ± 13 pmol/l, $p < 0.001$), and that leukocyte telomere length (LTL) was positively associated with freeT, independent of age and rare variants in telomerase genes ($P = 0.001$)²¹.

Approximately 25-60% of idiopathic PF patients have LTL below the 10th percentile²²⁻²⁶, suggesting telomere-mediated degenerative disease, and the proportion is much higher amongst familial cases with identified variants²⁴. We hypothesised that sex hormones protect against onset and progression of PF in such cases. Using UK Biobank (UKBB) data for IPF cases and

It is made available under a [CC-BY-NC-ND 4.0 International license](#).

controls we therefore investigated associations between blood LTL²⁷, sex hormone concentrations and IPF frequency and progression, using observational regression, survival analyses and MR methods to explore causality.

Methods

Study design and participants

UK Biobank (UKBB) contains observational and genetic data from ~500,000 adult volunteers aged 37-73 years, collected across the UK between 1/3/2006 and 31/3/2010²⁸. Self-reported menopause age, hormone replacement therapy (HRT), clinical follow-up via hospital record linkage and LTL (available for 474,000 participants²⁷) were analysed. A cohort of 379,708 unrelated participants of European descent was derived using ancestral principal components. This included 1133 IPF cases (defined by primary or secondary Hospital Episode Statistics (HES) ICD10 code of J84.1) and 378,575 controls. Details of the derivation with genotyping, and ethical approvals have been described previously¹⁰.

Observational Associations

Key demographics were compared using logistic regression models in Stata 16.1, adjusting throughout for confounders of sex hormone concentrations (age, ever having smoked, BMI, social deprivation, and fasting time where hormone levels were used), unless stated otherwise. The majority of females in UKBB are postmenopausal with oestrogen levels below the limit of detection, therefore associations with premature ovarian failure or primary ovarian insufficiency (defined here as menopause <40years), early menopause (age<45years), hysterectomy, bilateral oophorectomy and prescription/administration of HRT (oestrogen only and oestrogen + progesterone) were investigated using our derived variables^{29,30}. Associations between hormone change milestones (i.e., age of natural or surgical menopause) and age of PF diagnosis and all-cause death were also considered. HRT data were analysed by start date (pre- or post-2002) to avoid additional confounding following a 2002 publication on cancer risk which significantly impacted HRT prescribing practice in the UK³¹. Where specified, cases and controls were age-matched using 'calipmatch'³². Total testosterone, SHBG and serum albumin were measured at recruitment; bioavailable and free testosterone levels were derived using SHBG/albumin concentrations^{33,34} and inverse normalised for use in regression analyses³⁵.

Associations between IPF and its comorbidities were investigated with LTL as the exposure (using z-standardised natural logarithmic values of measured LTL²⁷). Associations of age-adjusted LTL with the key confounders known to affect sex hormone concentrations were computed, and links (adjusted for confounders) were examined between hormone-related medical conditions with LTL as the outcome. Only data from post-menopausal women (over age 55 at registration) were used to avoid selection bias due to the UKBB age profile.

Cox proportional hazards regression was used to explore onset of, and survival with, IPF in males and females separately, using LTL, sex hormone measurements and covariates. Start time for onset was UKBB registration until date of diagnosis, or from date of diagnosis until HES-recorded death. Survival was censored to the HES data cut-off date (27/03/2017). Proportional hazards assumptions were tested based on Schoenfeld residuals.

Mendelian Randomisation (MR)

MR is a statistical approach that uses natural genetic variation and its influence on phenotype to provide information about an exposure and an outcome³⁶. Since genetic variants are randomly assigned at conception, confounding and reverse causality are unlikely, making the genetic approach akin to a randomised control trial. Genetic variants robustly associated with sex hormone related outcomes (e.g. early menopause age <45 years) derived from genome wide association studies were used to create

It is made available under a [CC-BY-NC-ND 4.0 International license](#).

genetic instruments for early menopause in women³⁷, total and free testosterone concentrations in men³⁸ and SHBG concentrations in both³⁸ (see Supplement for details). The genetic risk score (GRS) for SHBG ('male SHBG cluster')³⁸ has been described previously as a genetic instrument with primary SHBG-increasing effects, and secondary divergent effects on total (higher) and bioavailable testosterone (lower), consistent with the known hormone-regulatory role of SHBG. For menopause age, two GRS were tested, the first containing 56 SNPs derived from meta-analyses independent of UKBB data³⁷ and the second containing 290 SNPs based on meta-analyses including UKBB data but with effect estimates generated in GWAS excluding UKBB³⁹, thus reducing the influence of Winner's curse⁴⁰. GRS for telomere length was previously reported¹⁰. Association between IPF prevalence, age of diagnosis, age of death and the GRS in the unrelated data sets of 204,736 females and 174,972 males was investigated using one-sample MR with logistic regression models. Ancestral principal components (as previously described⁴¹) were included as covariates in the analysis to control for residual population structure; results were adjusted for UKBB assessment centre and array chip. Where numbers were very low for MR (i.e. when investigating links with age of diagnosis [N=402] and age of death [N=199] amongst female IPF cases), the cohorts were divided into deciles and MR performed using the lowest and highest deciles, with corresponding sample sizes, n=31 and n=29 respectively (reduced by missing data).

Role of the funding source

Funders had no involvement in: study design; collection; analysis; data interpretation; report writing or the decision to submit the paper for publication. AD, KR, JT and CJS had access to the raw data. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

Results

Demographics

The well-established association between IPF incidence and male sex was replicated in this UKBB data set, after confounder adjustments: OR = 1.70 [95%CI: 1.50-1.93], $P < 1 \times 10^{-15}$. The association of male sex with death amongst the 1,133 unrelated IPF cases was also consistent with previous reports¹⁹: OR=1.58 [95%CI: 1.22-2.06], $P = 6.3 \times 10^{-4}$. Established associations were noted between IPF and a range of demographic and environmental variables such as age and smoking, plus newly reported strong associations with the two variables relating to leukocyte telomere length (LTL): adjusted T/S ratio and Z-adjusted T/S log (see Supplement), (Tables 1 & 2). A 1-SD decrease in LTL increased the odds of IPF by 1.58 (95%CI: 1.49-1.67), $P = 1 \times 10^{-57}$.

Demographic	Female IPF	Female Controls	OR	95% CI		P
N = 204,736 (%)	415 (36.6%)	204,321 (54.0%)				
Mean age at baseline (SD)	62.3 (6.1)	57.0 (7.9)	1.11	1.09	1.13	$< 1 \times 10^{-15}$
Mean age at diagnosis (SD)	65.6 (8.0)					
Townsend Deprivation Index (SD)	-0.83 (3.36)	-1.51 (2.94)	1.09	1.05	1.12	4.9×10^{-8}
Pollution, NO ₂ µg/m ³ (SD)	27.5 (8.8)	26.2 (7.4)	1.03	1.01	1.04	2.8×10^{-5}
Smoking status						
Never smoker (N, %)	159 (38.3%)	119,293 (58.4%)				
Former smoker (N, %)	190 (45.8%)	65,998 (32.3%)	1.97	1.59	2.43	3.4×10^{-10}
Current smoker (N, %)	59 (14.2%)	16,729 (8.2%)	3.13	2.32	4.22	9.9×10^{-14}
Missing (N, %)	7 (1.7%)	2,301 (1.1%)				
Body Mass Index (SD)	28.6 (5.7)	27.0 (5.1)	1.05	1.03	1.07	8.2×10^{-9}
Median income band (IQR)	2 (1-2)	2 (1-3)	0.65	0.58	0.74	1.1×10^{-11}
Mean FEV1 L (SD)	1.92 (0.48)	2.35 (0.53)	0.31	0.24	0.39	$< 1 \times 10^{-15}$
FEV1 percent predicted (SD)	68.8 (15.7)	77.3 (14.4)	0.97	0.96	0.98	$< 1 \times 10^{-15}$
Mean FVC L (SD)	2.56 (0.61)	3.08 (0.66)	0.36	0.30	0.44	$< 1 \times 10^{-15}$
FVC percent predicted (SD)	108 (23.5)	119 (21.3)	0.98	0.97	0.98	$< 1 \times 10^{-15}$
Metabolic Equivalent Task score (SD)	7.00 (1.30)	7.37 (0.92)	0.75	0.69	0.82	$< 1 \times 10^{-15}$
Adjusted T/S ratio (SD)	0.78 (0.13)	0.84 (0.13)	0.041	0.017	0.097	2.9×10^{-13}
Z-adjusted T/S log	-0.43 (1.08)	0.07 (0.99)	0.69	0.63	0.75	$< 1 \times 10^{-15}$
Participants deceased, N (%)	126 (30.4%)	5,066 (2.5%)	12.7	10.2	15.8	$< 1 \times 10^{-15}$

Table 1: Demographics for unrelated female idiopathic pulmonary fibrosis cases and controls of European ancestry in UK Biobank derived using logistic regression analyses. Odds ratios show the extent to which the exposures predicted prevalence of IPF, except for 'participants deceased', which shows the risk ratio for death with IPF. (Odds ratios and p values are adjusted for age, other than percent predicted spirometry which already includes adjustment). Income bands indicate average household income: (1) <£18000, (2) £18000-£30999, (3) £31000-51999, (4) £52000-£100000 and (5) >£100000.

Demographic	Male IPF	Male Controls	OR	95% CI		P
N = 174,972 (%)	718 (63.4%)	174,254 (46.0%)				
Mean age at baseline (SD)	63.1 (5.8)	57.1 (8.0)	1.14	1.12	1.14	$< 1 \times 10^{-15}$
Mean age at diagnosis (SD)	67.7 (7.2)					
Townsend Deprivation Index (SD)	-0.78 (3.31)	-1.44 (3.05)	1.09	1.06	1.11	1.5×10^{-13}
Pollution, NO ₂ µg/m ³ (SD)	27.1 (7.5)	26.3 (7.5)	1.02	1.01	1.03	5.1×10^{-5}
Smoking status						
Never smoker (N, %)	182 (25.4%)	84,584 (48.5%)				
Former smoker (N, %)	412 (57.4%)	67,823 (38.9%)	2.10	1.76	2.50	$< 1 \times 10^{-15}$
Current smoker (N, %)	108 (15.0%)	19,044 (10.9%)	2.83	2.23	3.60	$< 1 \times 10^{-15}$
Missing (N, %)	16 (2.2%)	2,803 (1.2%)				
Body Mass Index (SD)	28.5 (4.7)	27.8 (4.2)	1.04	1.03	1.06	1.1×10^{-5}
Median income band (IQR)	2 (1-3)	3 (2-4)	0.73	0.68	0.80	5.6×10^{-14}
Mean FEV1 L (SD)	2.60 (0.67)	3.24 (0.75)	0.44	0.38	0.49	$< 1 \times 10^{-15}$
FEV1 percent predicted (SD)	94.3 (22.1)	106.6 (20.7)	0.98	0.97	0.98	$< 1 \times 10^{-15}$
Mean FVC L (SD)	3.53 (0.81)	4.31 (0.92)	0.47	0.42	0.52	$< 1 \times 10^{-15}$
FVC percent predicted (SD)	107 (21.6)	120 (21.4)	0.97	0.97	0.98	$< 1 \times 10^{-15}$
Metabolic Equivalent Task score (SD)	7.11 (1.32)	7.45 (1.16)	0.79	0.74	0.84	2.6×10^{-13}
Adjusted T/S ratio (SD)	0.75 (0.13)	0.82 (0.13)	0.018	0.009	0.037	$< 1 \times 10^{-15}$
Z-adjusted T/S log	-0.74 (1.19)	-0.11 (0.99)	0.63	0.58	0.67	$< 1 \times 10^{-15}$
Participants deceased, N (%)	313 (43.6%)	7,838 (4.5%)	11.4	9.7	13.2	$< 1 \times 10^{-15}$

Tables 2: Demographics for unrelated male idiopathic pulmonary fibrosis cases and controls of European ancestry in UK Biobank derived using logistic regression analyses. Odds ratios show the extent to which the exposures predicted prevalence of IPF, except for 'participants deceased', which shows the risk ratio for death with IPF. (Odds ratios and p values are adjusted for age, other than percent predicted spirometry which already includes adjustment). Income bands indicate average household income: (1) <£18000, (2) £18000-£30999, (3) £31000-51999, (4) £52000-£100000 and (5) >£100000.

Observational associations of IPF and its comorbidities with measured LTL

Building on previously reported findings for multiple other diseases in both sexes together^{42,43}, IPF prevalence showed strong associations with LTL for both males (OR=0.63, $P=2.2 \times 10^{-38}$) and females (OR=0.69, $P=8.9 \times 10^{-16}$), with odds ratios much lower than those for other reported associated comorbidities of IPF such as coronary artery disease and type II diabetes (Figure 1, Supplement:S1). A similar, slightly weaker association was also found for a broader clinical definition of interstitial lung disease (ILD). Interestingly, the age-adjusted associations differed between sexes for some of these outcomes, including death.

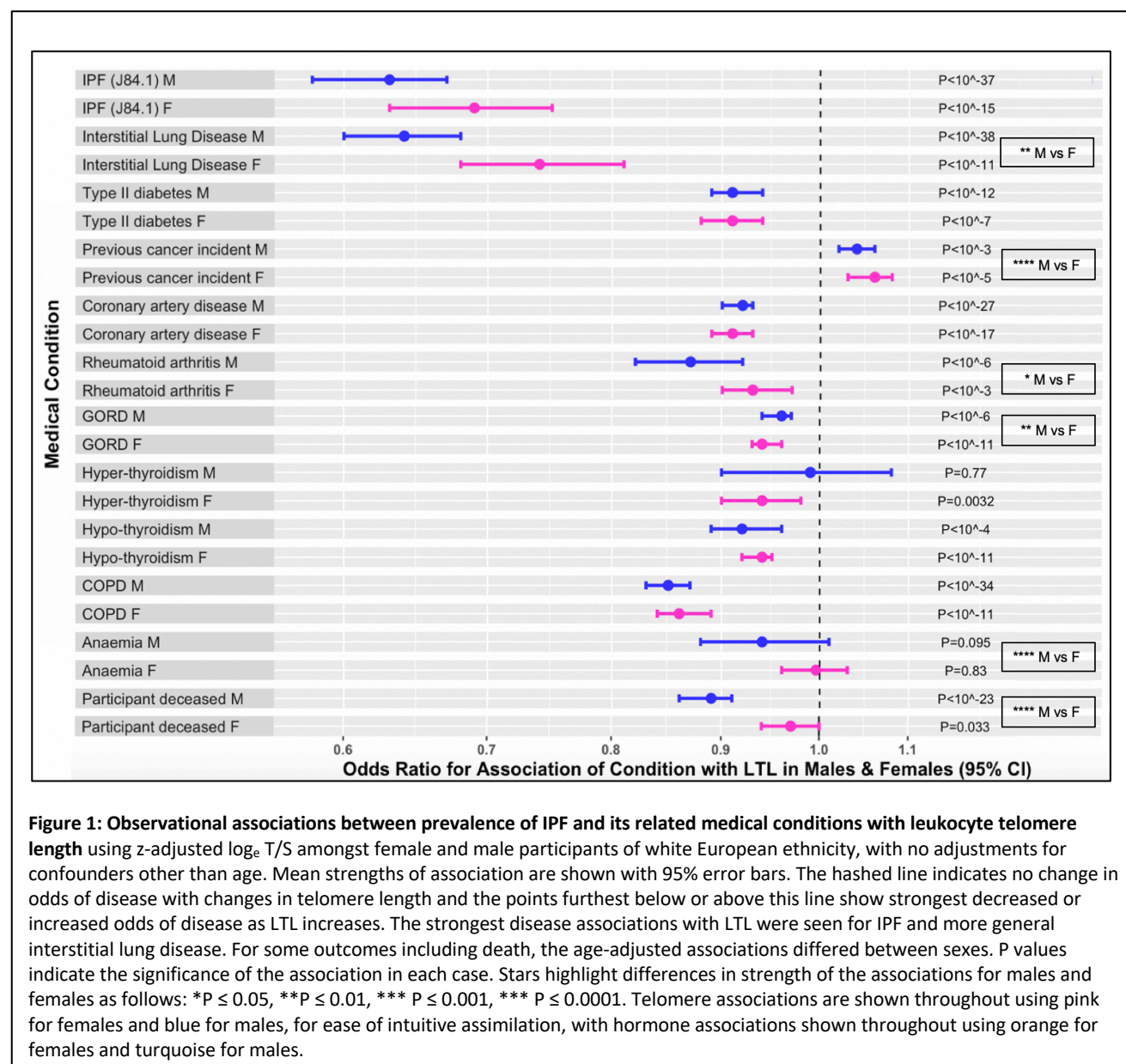


Figure 1: Observational associations between prevalence of IPF and its related medical conditions with leukocyte telomere length using z-adjusted \log_e T/S amongst female and male participants of white European ethnicity, with no adjustments for confounders other than age. Mean strengths of association are shown with 95% error bars. The hashed line indicates no change in odds of disease with changes in telomere length and the points furthest below or above this line show strongest decreased or increased odds of disease as LTL increases. The strongest disease associations with LTL were seen for IPF and more general interstitial lung disease. For some outcomes including death, the age-adjusted associations differed between sexes. P values indicate the significance of the association in each case. Stars highlight differences in strength of the associations for males and females as follows: * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. Telomere associations are shown throughout using pink for females and blue for males, for ease of intuitive assimilation, with hormone associations shown throughout using orange for females and turquoise for males.

Association between factors affecting sex hormones and LTL

Associations were found between shorter LTL and the potential confounders we identified as affecting sex hormone levels (smoking, BMI, social deprivation and fasting duration) (Figure 2, Supplement: S2), with large effect sizes for current and male smokers. Importantly for this study, and as reported elsewhere²⁷, the negative effect of age on LTL for females over the age of

55 (i.e. post-menopause) is greater than for younger women and similar to that of males whereas for younger women it is significantly better than for men.

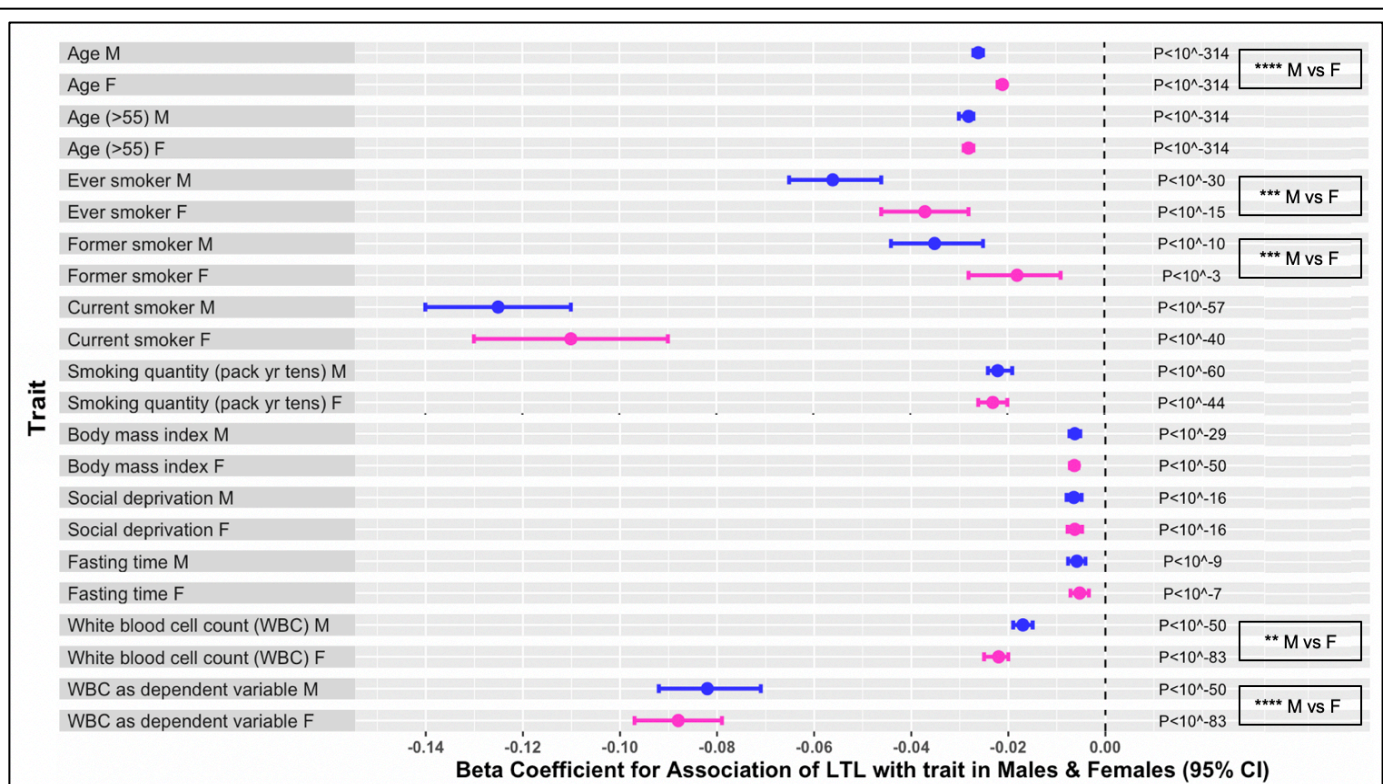


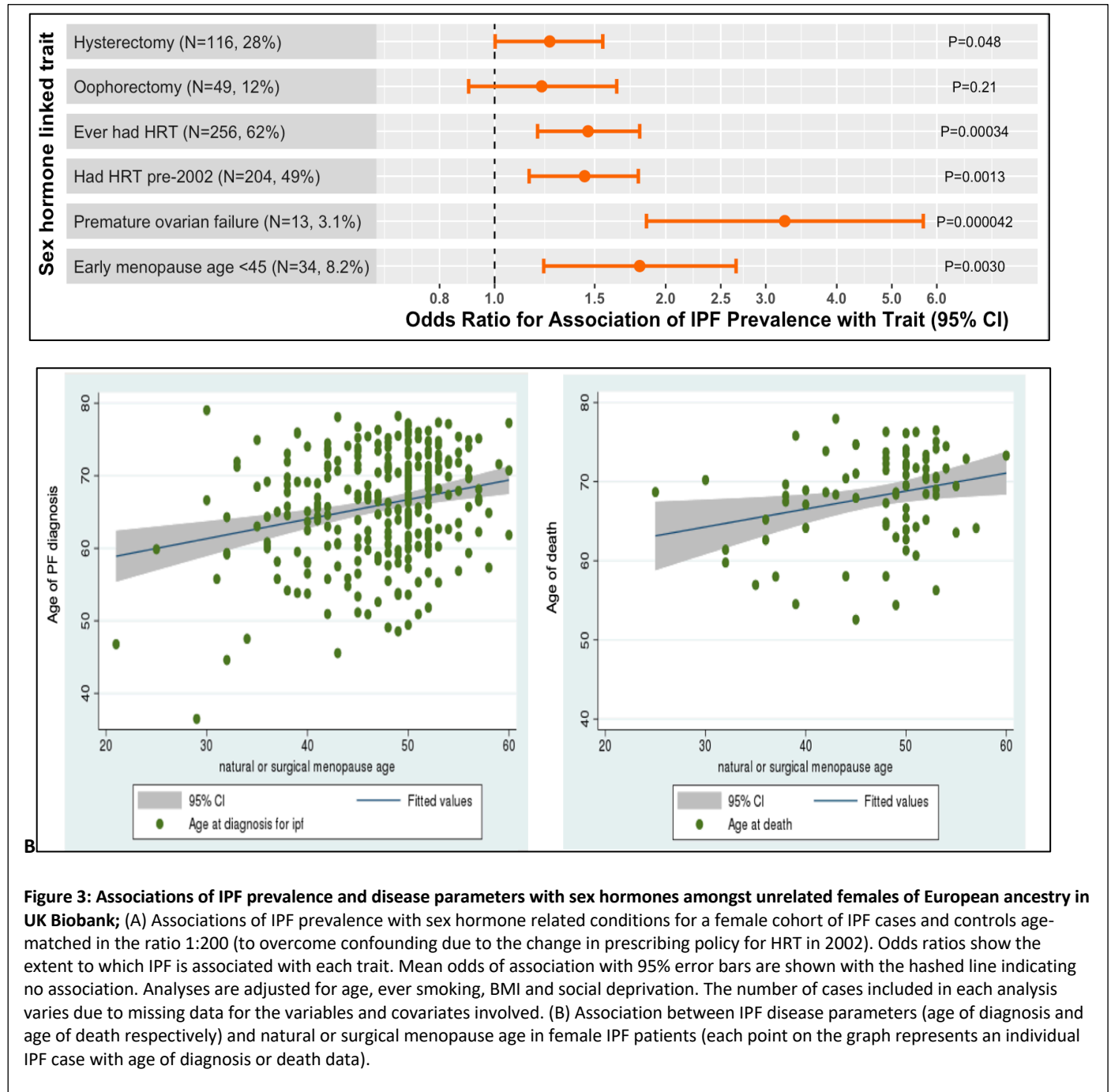
Figure 2: Observational associations of leukocyte telomere length using z-adjusted log_e T/S with key characteristics, in particular those known to be associated with sex hormone levels, amongst female and male participants of white European ethnicity, with no adjustments for confounders other than age. Mean strengths of association are shown with 95% error bars. The hashed line indicates no change in mean telomere length with the trait and the points furthest below this line show strongest odds of decreased LTL as the trait increases. Large effect sizes are seen for current and male smokers and the impact of age is smaller for pre-menopausal women than for older women and for men. Stars highlight differences in the means between males and females as follows: * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. White blood cell count was included here for consideration of its role as a possible confounder or dependent variable in telomere biology disorders (although the LTL measurement is independent of cell count).

Observational associations between PF and sex hormones

Amongst female IPF cases and age-matched controls (match ratio 1:200), IPF prevalence was associated with conditions of sex hormone depletion, after adjusting for the key confounders for menopause age³⁰ (Figure 3A, Supplement: S3a). The strongest association was for premature ovarian failure (menopause <40yrs); OR=3.24 [95%CI:1.85-5.68], $P=4.2 \times 10^{-5}$. IPF prevalence was also negatively associated with age of natural menopause: OR=0.95 [95%CI:0.92-0.97], $P=5.9 \times 10^{-5}$. Mean age of natural menopause was lower amongst IPF cases than age-matched controls: 48.9 (± 5.6 SD) years for IPF cases (N=213) vs 50.2 (± 4.5 SD) years for controls, β coefficient = -1.33 [95%CI: -1.94, -0.72], $P=2.0 \times 10^{-5}$. The mean gap between age of natural/surgical menopause and age of IPF diagnosis was 18.7 (± 8.8 SD) years, with only 5 cases being diagnosed prior to menopause. There was a significant association between natural or surgical menopause age and age of diagnosis for N=264 IPF cases (reduced by missing covariate data): β coefficient = 0.27 [95%CI:0.14-0.41], $P=8.9 \times 10^{-5}$ (Figure 3B). Similarly, there was an association

It is made available under a [CC-BY-NC-ND 4.0 International license](#).

between menopause age and age of death for N=80 cases, with on average a 10-year increase in menopause age corresponding to a 2.3 year longer life: β coefficient =0.22 [95%CI:0.024-0.41], P=0.028.



For male participants who later developed IPF, mean fractions of bioavailable testosterone and free testosterone were within normal range at registration but were lower than for age-matched controls (match ratio 1:100): bioavailable testosterone for cases was 39.6% (\pm 9.6 SD) compared with 42.3% (\pm 9.7 SD) for controls (β = -3.04, $P=3 \times 10^{-16}$) and free testosterone was 1.66% (\pm 0.40 SD) compared with 1.73% (\pm 0.39 SD) for controls (β = -0.081, $P=5 \times 10^{-8}$). Mean haemoglobin concentration for existing male IPF cases at registration was within normal range, at 14.5g/dL (\pm 1.6 SD).

It is made available under a [CC-BY-NC-ND 4.0 International license](#).

SHBG controls sex hormone bioavailability, with higher concentrations reducing the bioavailable and free testosterone *fractions* and raising the total testosterone concentration (Figure 4A-C, Supplement S5). After covariate adjustments, IPF prevalence in males was associated positively with SHBG, OR=1.35 [95%CI: 1.23-1.47], $P=2.6 \times 10^{-11}$ per 1SD higher concentration at registration. IPF was associated negatively with bioavailable testosterone percentage: OR=0.69 [95%CI: 0.63-0.76], $P=5.6 \times 10^{-16}$ per 1SD increase (Supplement: S5), even though the majority of cases (612/718 males) were incident, with mean time to diagnosis = 4.99 years \pm 2.32 SD. A weak association was also found amongst females: OR=0.85 [95%CI: 0.73-0.97], $P=0.02$. further observational associations between IPF and hormone level parameters are reported in the supplement.

In summary, the data shows consistent multi-faceted evidence of PF prevalence linked with reduced sex hormone concentrations.

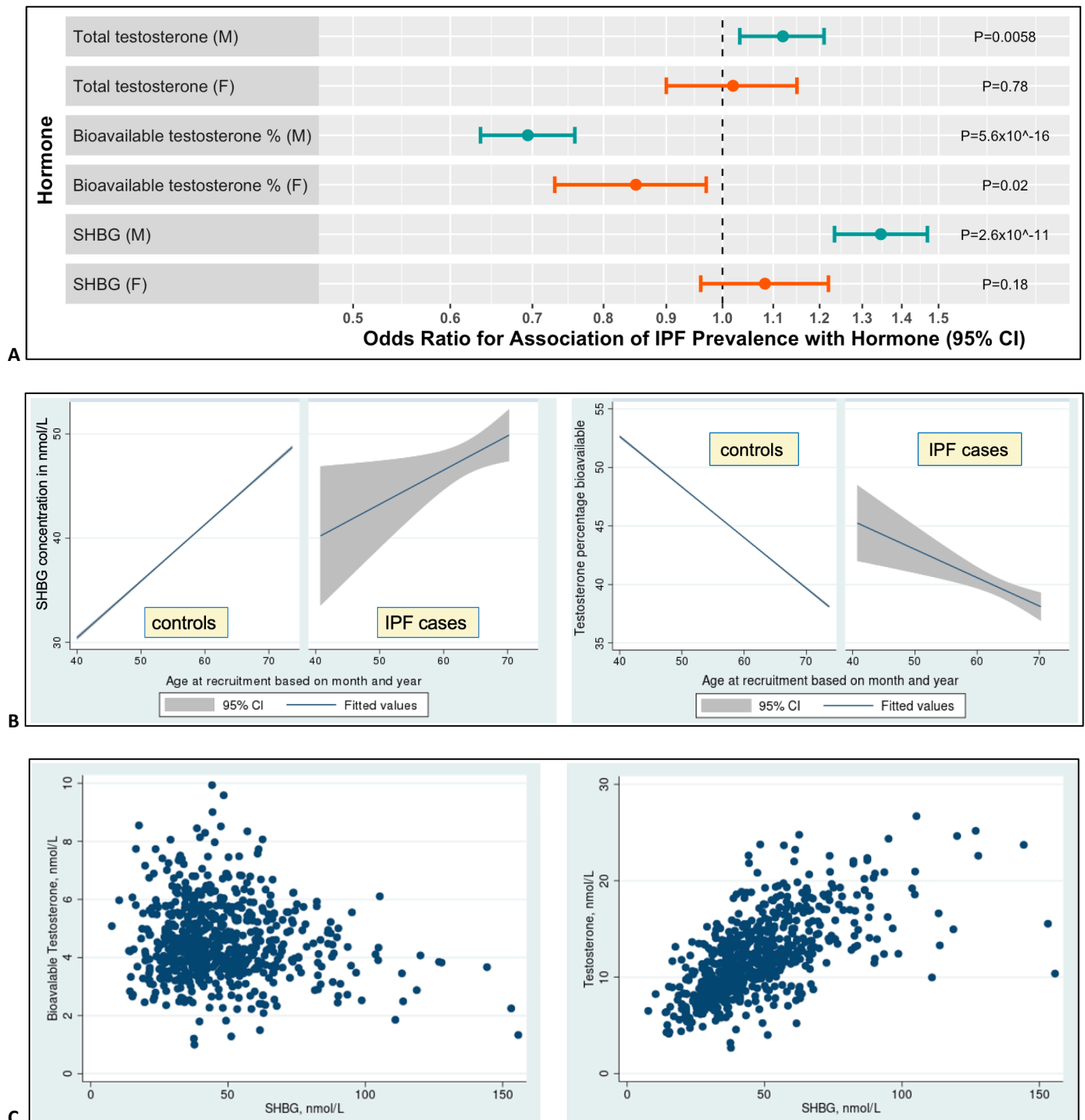
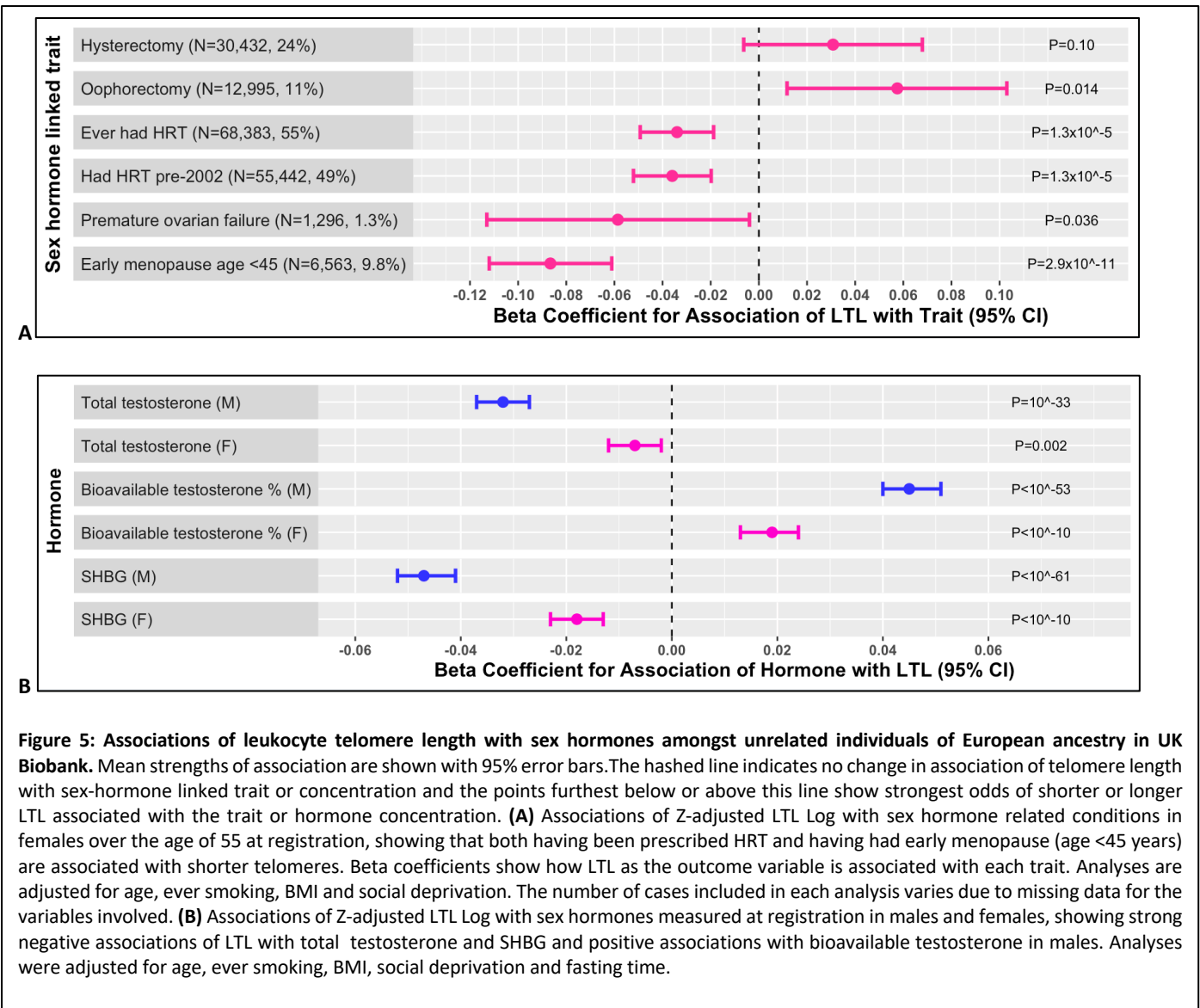


Figure 4: Associations of IPF prevalence with sex hormones amongst unrelated females and males of European ancestry in UK Biobank; (A) Odds of IPF prevalence amongst male and female IPF cases and controls for a 1SD (standard deviation) increase in sex hormone levels. Analyses are adjusted for age, ever smoking, BMI, social deprivation and fasting time. (B) Mean and 95% CI concentrations of SHBG and of bioavailable testosterone with age at recruitment for male controls and IPF cases (for controls, the 95% lines are very similar to the mean due to high numbers so not visible in the plot). (C) Scatter plots showing how bioavailable and total testosterone concentrations vary with SHBG in IPF (each point on the graph represents an individual IPF case). Sex hormone associations are shown throughout using orange for females and turquoise for males.

LTL and sex hormones

The above conditions relating to sex hormone levels were also associated with shorter telomeres (Figure 5A, Supplement: S6A), with a striking association with early menopause: $\beta = -0.087$ [95%CI:0.112-0.61], $P=2.9 \times 10^{-11}$. There was no observed association of hysterectomy and oophorectomy with *shorter* telomeres, likely because the indications for surgery are diverse but typically included as part of cancer treatment. LTL was also positively associated with natural menopause age: $\beta = 0.0080$ [95%CI:0.0065-0.0094], $P=6.6 \times 10^{-27}$.



Mirroring the associations with IPF, LTL was associated negatively with total testosterone and SHBG but positively with bioavailable testosterone, again with considerably stronger associations in males than females (Figure 5B, Supplement: S6b).

Sex hormones and disease progression

Multiple cox proportional hazard regression analyses for PF onset and survival showed lower hazard ratios linked with increasing LTL and sex hormone levels. Slower disease progression was linked with both later menopause and having received HRT for females and with higher bioavailable testosterone concentrations for males (see supplemental Results). This suggests that for patients with PF, restored hormone levels could be important for survival.

Genetic assessment of causality

Evidence of a causal link between our SHBG GRS, 'male SHBG cluster' and telomere length was found amongst all UKBB participants: $\beta = -0.00085$ [95%CI: -0.0012, -0.0052], $P = 6.5 \times 10^{-7}$ (N=365,075; adjusting for principal components, test centre, chip, fasting time and age), suggesting that higher SHBG leads to shorter telomeres and could therefore be implicated in PF causality. This GRS has been described previously and represents a genetic instrument with primary SHBG-increasing effects, and secondary divergent effects on total (higher) and bioavailable testosterone (lower) that are consistent with the known hormone-regulatory role of SHBG.

Using our telomere length GRS¹⁰ (which is strongly associated with measured LTL: $R^2 = 0.009$, $F = 3411$), there also was evidence suggesting that SHBG levels (single inverse normalised) are affected by shorter telomeres in men: $\beta = 0.011$ [95%CI: 0.008-0.014], $P = 7.9 \times 10^{-12}$. Amongst women, there is weaker evidence of a link: $\beta = 0.0038$ [95%CI: 0.009-0.0067], $P = 0.010$. Thus, in men, short telomeres appear to cause higher SHBG levels which would reduce the bioavailable testosterone further. This finding supports the possibility of a detrimental feedback loop for high SHBG and short telomeres amongst males with IPF. If a more stringent p value of $p < 0.001$ is used to allow for multiple testing with 6 different hormone genetic risk scores, the above results in males remain robust. Additional MR analyses hinted at causality behind our reported observations but were limited by IPF case numbers that are too low for this type of analysis (see Supplement).

Thus, our genetic analysis provides evidence suggesting a bi-directional causal role in IPF for increased SHBG levels and short telomeres in males and some indications of a causal role for short telomeres and sex hormone insufficiency in females. Future large-scale genetic studies will enable further analyses of these links.

Discussion

Observational analyses in UKBB data have shown evidence of multiple consistent associations between pulmonary fibrosis and lower sex hormone levels in males and females. These findings were underpinned by links between PF and short LTL which were then mirrored in relationships between LTL and hormone levels. Genetic causal analyses provided evidence of a causal interaction between SHBG and telomere length, most important in males where it appears to be bi-directional. Taken together, these findings support our hypothesis that high-normal sex hormone concentrations could be protective in PF, whereas low-normal sex hormone concentrations could be detrimental.

Recent findings have reported a strong observational association of longer telomeres with menopause age and both observational and MR associations between longer telomeres and lower SHBG and testosterone (males)⁴². Our results build on

It is made available under a [CC-BY-NC-ND 4.0 International license](#).

these, by highlighting the strength of association with LTL of IPF and ILD and through carrying out separate analyses in males and females to elicit important gender differences.

Reports suggest that sex hormones play a regulatory role in expression of TERT, with oestrogen receptors appearing to act directly on the *TERT* gene promoter to upregulate telomerase^{44,45,46} and androgens acting indirectly via a range of mechanisms including (but not limited to) aromatisation to oestrogen^{47-49,50,51}. Of particular interest, in cell culture of peripheral blood lymphocytes from patients heterozygous for telomerase mutations, low baseline telomerase levels were restored to normal by androgen exposure⁴⁹. Amongst healthy women, long-term hormone replacement therapy (HRT) correlates with longer telomeres⁵².

Evidence for the potential therapeutic role of sex hormones in PF exists in the treatment of 27 TBD patients (15 female, median age 41 years and age range 17-66), with the synthetic androgen danazol⁵³. Whilst caution in interpretation is required as danazol is a weak androgen, weak progestogen, weak antigonadotrophin, weak steroidogenesis inhibitor and a functional antioestrogen, recruitment was halted early because 11/12 patients had an unanticipated gain in telomere length at 24 months. PF scores based on CT⁵⁴ and lung function DLCO measurements were stable during the 2 years of treatment for all patients but one. Due to side effects, several trials at lower doses are now underway⁵⁵⁻⁵⁷. In a dramatic recent case study, a 42 year old female PF patient due for lung transplant recovered after 18 months on danazol⁵⁸.

However, adopting a therapeutic approach to supplement sex hormone concentrations above physiological requirements requires key safety considerations and monitoring implications. Supraphysiological testosterone administration stimulates erythrocytosis and so causes secondary polycythaemia and hyperviscosity, which can increase the risk of cardio- and cerebrovascular events. Women beyond the age of natural menopause are only recommended to take HRT for 5 years to minimise the risk of stimulated breast cancer and increased thromboembolic events. However, as sex steroid concentrations were lower than age-matched controls, supplementing to higher concentrations within the normal range but not above could yield disease benefit in IPF but avoid at least some of the side effects. Similarly, upregulation of telomerase carries risks. Although overt TERT expression is insufficient to cause malignant transformation, its presence can promote unchecked proliferation; over-expression of telomerase is detected in around 90% of human cancers¹⁵. Shorter telomeres are generally associated with mortality^{42,43}, yet we found this to be the case for men only (Figure 2), probably because a higher proportion of female deaths are caused by cancer, which is associated with longer telomeres (in UKBB, 68% of deaths amongst the top 20 causes for females were cancer-related whereas for males, this figure was 47%⁵⁹).

Our study is limited by the lack of extensive clinical characterisation of cases (although these have shown good consistency with replication cohorts previously¹⁰), and by the healthy volunteer effect amongst research participants. The MR studies amongst IPF cases are currently underpowered and replication in large clinical cohorts would be valuable. Some of the key variables (e.g. menopause age) are self-reported and so may be impacted by recall bias.

In summary, these findings provide evidence that lower sex hormone concentrations are related to PF and we speculate that treatment to supplement sex hormone concentrations higher within the normal reference range but not above it, may delay or prevent onset of PF amongst at risk individuals (such as relatives of patients with familial PF caused by germline telomere defects) and slow disease progression amongst patients with short telomeres. If correct, this may offer a relatively low-cost (particularly when compared to novel agents for IPF), available, and well-tolerated treatment which could be explored in a future, randomised clinical trial powered for safety and disease outcomes in men and women.

Author Contributions

AD conceived the project, carried out the analyses, interpreted the results and drafted the manuscript; AM, KSR, RNB, ARW and SM contributed to the statistical analyses, MAG, JKP and AMR provided clinical guidance and interpretation; HA and JC provided patient representation; KL, and MAL provided supervisory support to AD; JT and CJS were the principal investigators for the study, with oversight of study design, analysis and interpretation, plus administration of access to UK Biobank. All authors contributed to drafting and revision of the manuscript. All authors read and approved of the final manuscript.

Acknowledgements

This work was supported in part by grant MR/N0137941/1 for the GW4 BIOMED MRC DTP, awarded to the Universities of Bath, Bristol, Cardiff and Exeter from the Medical Research Council (MRC)/UKRI. KSR is supported by Cancer Research UK [grant number C18281/A29019]. JKP is supported by the UKRI Expanding Excellence in England award. JT is supported by an Academy of Medical Sciences (AMS) Springboard award, which is supported by the AMS, the Wellcome Trust, GCRF, the Government Department of Business, Energy and Industrial strategy, the British Heart Foundation and Diabetes UK [SBF004\1079]. CJS is supported by MRC project grants (MR/V002538/1 and MR/S002626/1). This research has been conducted using the UK Biobank Resource (applications 9072 and 44046).

Declaration of interests

MAG has received support to attend conferences and professional fees from Roche and Boehringer-Ingelheim, outside of the submitted work. LVW has received grants from GSK, outside the submitted work. AD, KSR, AMR, HA, JC, KL, MAL, JKP, JT and CJS have no declaration of interest.

References

1. Marshall DC, Salciccioli JD, Shea BS, Akuthota P. Trends in mortality from idiopathic pulmonary fibrosis in the European Union: an observational study of the WHO mortality database from 2001–2013. *European Respiratory Journal* 2018; **51**(1): 1701603.
2. Snell N, Strachan D, Hubbard R, Gibson J, Maher T, Jarrold I. P272 Epidemiology of idiopathic pulmonary fibrosis in the uk: findings from the british lung foundation's 'respiratory health of the nation' project. *Thorax* 2016; **71**(Suppl 3): A236-A.
3. Manni A, Pardridge WM, Cefalu W, et al. Bioavailability of albumin-bound testosterone. *J Clin Endocrinol Metab* 1985; **61**(4): 705-10.
4. Goldman AL, Bhasin S, Wu FCW, Krishna M, Matsumoto AM, Jasuja R. A Reappraisal of Testosterone's Binding in Circulation: Physiological and Clinical Implications. *Endocrine Reviews* 2017; **38**(4): 302-24.
5. Armanios MY, Chen JJ, Cogan JD, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007; **356**(13): 1317-26.
6. Tsakiri KD, Cronkhite JT, Kuan PJ, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proceedings of the National Academy of Sciences of the United States of America* 2007; **104**(18): 7552-7.
7. Blackburn EH, Greider CW, Szostak JW. Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nat Med* 2006; **12**(10): 1133-8.
8. Dai J, Cai H, Li H, et al. Association between telomere length and survival in patients with idiopathic pulmonary fibrosis. *Respirology* 2015; **20**(6): 947-52.
9. Stuart BD, Lee JS, Kozlitina J, et al. Effect of telomere length on survival in idiopathic pulmonary fibrosis: an observational study with independent validation. *The Lancet Respiratory medicine* 2014; **2**(7): 557-65.
10. Duckworth A, Gibbons MA, Allen RJ, et al. Telomere length and risk of idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease: a mendelian randomisation study. *The Lancet Respiratory Medicine* 2020.
11. Calado RT, Young NS. Telomere maintenance and human bone marrow failure. *Blood* 2008; **111**(9): 4446-55.
12. Khincha PP, Wentzensen IM, Giri N, Alter BP, Savage SA. Response to androgen therapy in patients with dyskeratosis congenita. *Br J Haematol* 2014; **165**(3): 349-57.
13. Kirschner M, Vieri M, Kricheldorf K, et al. Androgen derivatives improve blood counts and elongate telomere length in adult cryptic dyskeratosis congenita. *Br J Haematol* 2021; **193**(3): 669-73.
14. Khincha PP, Bertuch AA, Gadalla SM, Giri N, Alter BP, Savage SA. Similar telomere attrition rates in androgen-treated and untreated patients with dyskeratosis congenita. *Blood advances* 2018; **2**(11): 1243-9.
15. Shay JW, Bacchetti S. A survey of telomerase activity in human cancer. *European Journal of Cancer* 1997; **33**(5): 787-91.
16. Garate-Carrillo A, Gonzalez J, Ceballos G, Ramirez-Sanchez I, Villarreal F. Sex related differences in the pathogenesis of organ fibrosis. *Transl Res* 2020; **222**: 41-55.
17. Cheung AS, Yeap BB, Hoermann R, Hui J, Beilby JP, Grossmann M. Effects of androgen deprivation therapy on telomere length. *Clin Endocrinol (Oxf)* 2017; **87**(4): 381-5.
18. Gardner M, Bann D, Wiley L, et al. Gender and telomere length: systematic review and meta-analysis. *Exp Gerontol* 2014; **51**: 15-27.
19. Foundation BL. The battle for breath: the impact of lung disease in the UK, 2016.

20. Han MK, Murray S, Fell CD, et al. Sex differences in physiological progression of idiopathic pulmonary fibrosis. *The European respiratory journal* 2008; **31**(6): 1183-8.
21. Fang C, Huang H, Zhang Q, et al. Relation between sex hormones and leucocyte telomere length in men with idiopathic pulmonary fibrosis. *Respirology*; n/a(n/a).
22. Hunninghake GM, Quesada-Arias LD, Carmichael NE, et al. Interstitial Lung Disease in Relatives of Patients with Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine* 2020; **201**(10): 1240-8.
23. Cronkhite JT, Xing C, Raghu G, et al. Telomere Shortening in Familial and Sporadic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine* 2008; **178**(7): 729-37.
24. Newton CA, Batra K, Torrealba J, et al. Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive. *Eur Respir J* 2016; **48**(6): 1710-20.
25. Newton CA, Oldham JM, Ley B, et al. Telomere length and genetic variant associations with interstitial lung disease progression and survival. *European Respiratory Journal* 2019; **53**(4): 1801641.
26. Newton CA, Zhang D, Oldham JM, et al. Telomere Length and Use of Immunosuppressive Medications in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2019; **200**(3): 336-47.
27. Codd V, Denniff M, Swinfield C, et al. A major population resource of 474,074 participants in UK Biobank to investigate determinants and biomedical consequences of leukocyte telomere length. *medRxiv* 2021: 2021.03.18.21253457.
28. Sudlow C, Gallacher J, Allen N, 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 medicine* 2015; **12**(3): e1001779-e.
29. Ruth KS, Beaumont RN, Tyrrell J, et al. Genetic evidence that lower circulating FSH levels lengthen menstrual cycle, increase age at menopause and impact female reproductive health. *Hum Reprod* 2016; **31**(2): 473-81.
30. Ruth KS, Perry JRB, Henley WE, Melzer D, Weedon MN, Murray A. Events in Early Life are Associated with Female Reproductive Ageing: A UK Biobank Study. *Sci Rep* 2016; **6**: 24710-.
31. Investigators WGfWShI. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women Principal Results From the Women's Health Initiative Randomized Controlled Trial. *JAMA* 2002; **288**(3): 321-33.
32. Stepner M, Garland A. CALIPMATCH: Stata module for caliper matching without replacement. 2017.
33. Vermeulen A, Verdonck L, Kaufman JM. A Critical Evaluation of Simple Methods for the Estimation of Free Testosterone in Serum. *The Journal of Clinical Endocrinology & Metabolism* 1999; **84**(10): 3666-72.
34. Chung MC, Gombar S, Shi RZ. Implementation of Automated Calculation of Free and Bioavailable Testosterone in Epic Beaker Laboratory Information System. *J Pathol Inform* 2017; **8**: 28-.
35. Laurent MR, Hammond GL, Blokland M, et al. Sex hormone-binding globulin regulation of androgen bioactivity in vivo: validation of the free hormone hypothesis. *Sci Rep* 2016; **6**(1): 35539.
36. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014; **23**(R1): R89-98.
37. Day FR, Ruth KS, Thompson DJ, et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nature genetics* 2015; **47**(11): 1294-303.
38. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine* 2020; **26**(2): 252-8.
39. Ruth KS, Day FR, Hussain J, et al. Genetic insights into biological mechanisms governing human ovarian ageing. *Nature* 2021; **596**(7872): 393-7.

40. Kraft P. Curses—Winner's and Otherwise—in Genetic Epidemiology. *Epidemiology* 2008; **19**(5).
41. Biobank U. UK Biobank Genetic Data: MRC-IEU Quality Control, version 1. 2017.
https://data.bris.ac.uk/datasets/3074krb6t2frj29yh2b03x3wxj/UK%20Biobank%20Genetic%20Data_MRC%20IEU%20Quality%20Control%20version%201.pdf (accessed 3/11/18 2018).
42. Codd V, Wang Q, Allara E, et al. Polygenic basis and biomedical consequences of telomere length variation. *Nat Genet* 2021; **53**(10): 1425-33.
43. Schneider CV, Schneider KM, Teumer A, et al. Association of Telomere Length With Risk of Disease and Mortality. *JAMA Internal Medicine* 2022; **182**(3): 291-300.
44. Kyo S, Takakura M, Kanaya T, et al. Estrogen Activates Telomerase. *Cancer research* 1999; **59**(23): 5917-21.
45. Nanni S, Narducci M, Della Pietra L, et al. Signaling through estrogen receptors modulates telomerase activity in human prostate cancer. *The Journal of Clinical Investigation* 2002; **110**(2): 219-27.
46. Sato R, Maesawa C, Fujisawa K, et al. Prevention of critical telomere shortening by oestradiol in human normal hepatic cultured cells and carbon tetrachloride induced rat liver fibrosis. *Gut* 2004; **53**(7): 1001-9.
47. Guo C, Armbruster BN, Price DT, Counter CM. In vivo regulation of hTERT expression and telomerase activity by androgen. *The Journal of urology* 2003; **170**(2 Pt 1): 615-8.
48. Nourbakhsh M, Golestani A, Zahrai M, Modarressi MH, Malekpour Z, Karami-Tehrani F. Androgens stimulate telomerase expression, activity and phosphorylation in ovarian adenocarcinoma cells. *Molecular and Cellular Endocrinology* 2010; **330**(1): 10-6.
49. Calado RT, Yewdell WT, Wilkerson KL, et al. Sex hormones, acting on the TERT gene, increase telomerase activity in human primary hematopoietic cells. *Blood* 2009; **114**(11): 2236-43.
50. Yeap BB, Knuiman MW, Divitini ML, et al. Epidemiological and Mendelian Randomization Studies of Dihydrotestosterone and Estradiol and Leukocyte Telomere Length in Men. *J Clin Endocrinol Metab* 2016; **101**(3): 1299-306.
51. Li CL, Li CY, Lin YY, et al. Androgen Receptor Enhances Hepatic Telomerase Reverse Transcriptase Gene Transcription After Hepatitis B Virus Integration or Point Mutation in Promoter Region. *Hepatology (Baltimore, Md)* 2019; **69**(2): 498-512.
52. Lee D-C, Im J-A, Kim J-H, Lee H-R, Shim J-Y. Effect of Long-Term Hormone Therapy on Telomere Length in Postmenopausal Women. *Yonsei Med J* 2005; **46**(4): 471-9.
53. Townsley DM, Dumitriu B, Liu D, et al. Danazol Treatment for Telomere Diseases. *The New England journal of medicine* 2016; **374**(20): 1922-31.
54. Rosas IO, Yao J, Avila NA, Chow CK, Gahl WA, Gochuico BR. Automated Quantification of High-Resolution CT Scan Findings in Individuals at Risk for Pulmonary Fibrosis. *CHEST* 2011; **140**(6): 1590-7.
55. Low-Dose Danazol for the Treatment of Telomere Related Diseases.
<https://ClinicalTrials.gov/show/NCT03312400>.
56. Mackintosh JA, Pietsch M, Lutzky V, et al. TELO-SCOPE study: a randomised, double-blind, placebo-controlled, phase 2 trial of danazol for short telomere related pulmonary fibrosis. *BMJ Open Respir Res* 2021; **8**(1).
57. Evaluating Efficacy and Safety of Danazol in Severe Hematologic or Pulmonary Disease Related to Telomeropathy (ANDROTELO). 2018. <https://clinicaltrials.gov/ct2/show/NCT03710356>.
58. Chambers DC, Lutzky VP, Apte SH, Godbolt D, Feenstra J, Mackintosh J. Successful treatment of telomeropathy-related interstitial lung disease with immunosuppression and danazol. *Respirology case reports* 2020; **8**(6): e00607.
59. UK Biobank Death Summary Report. 2021.

Tables

Demographic	Female IPF	Female Controls	OR	95% CI		P
N = 204,736 (%)	415 (36.6%)	204,321 (54.0%)				
Mean age at baseline (SD)	62.3 (6.1)	57.0 (7.9)	1.11	1.09	1.13	<1x10 ⁻¹⁵
Mean age at diagnosis (SD)	65.6 (8.0)					
Townsend Deprivation Index (SD)	-0.83 (3.36)	-1.51 (2.94)	1.09	1.05	1.12	4.9x10 ⁻⁸
Pollution, NO ₂ µg/m ³ (SD)	27.5 (8.8)	26.2 (7.4)	1.03	1.01	1.04	2.8x10 ⁻⁵
Smoking status						
Never smoker (N, %)	159 (38.3%)	119,293 (58.4%)				
Former smoker (N, %)	190 (45.8%)	65,998 (32.3%)	1.97	1.59	2.43	3.4x10 ⁻¹⁰
Current smoker (N, %)	59 (14.2%)	16,729 (8.2%)	3.13	2.32	4.22	9.9x10 ⁻¹⁴
Missing (N, %)	7 (1.7%)	2,301 (1.1%)				
Body Mass Index (SD)	28.6 (5.7)	27.0 (5.1)	1.05	1.03	1.07	8.2x10 ⁻⁹
Median income band (IQR)	2 (1-2)	2 (1-3)	0.65	0.58	0.74	1.1x10 ⁻¹¹
Mean FEV1 L (SD)	1.92 (0.48)	2.35 (0.53)	0.31	0.24	0.39	<1x10 ⁻¹⁵
FEV1 percent predicted (SD)	68.8 (15.7)	77.3 (14.4)	0.97	0.96	0.98	<1x10 ⁻¹⁵
Mean FVC L (SD)	2.56 (0.61)	3.08 (0.66)	0.36	0.30	0.44	<1x10 ⁻¹⁵
FVC percent predicted (SD)	108 (23.5)	119 (21.3)	0.98	0.97	0.98	<1x10 ⁻¹⁵
Metabolic Equivalent Task score (SD)	7.00 (1.30)	7.37 (0.92)	0.75	0.69	0.82	<1x10 ⁻¹⁵
Adjusted T/S ratio (SD)	0.78 (0.13)	0.84 (0.13)	0.041	0.017	0.097	2.9x10 ⁻¹³
Z-adjusted T/S log	-0.43 (1.08)	0.07 (0.99)	0.69	0.63	0.75	<1x10 ⁻¹⁵
Participants deceased, N (%)	126 (30.4%)	5,066 (2.5%)	12.7	10.2	15.8	<1x10 ⁻¹⁵

Table 1: Demographics for unrelated female idiopathic pulmonary fibrosis cases and controls of European ancestry in UK Biobank derived using logistic regression analyses. Odds ratios show the extent to which the exposures predicted prevalence of IPF, except for 'participants deceased', which shows the risk ratio for death with IPF. (Odds ratios and p values are adjusted for age, other than percent predicted spirometry which already includes adjustment). Income bands indicate average household income: (1) <£18000, (2) £18000-£30999, (3) £31000-51999, (4) £52000-£100000 and (5)>£100000.

Demographic	Male IPF	Male Controls	OR	95% CI		P
N = 174,972 (%)	718 (63.4%)	174,254 (46.0%)				
Mean age at baseline (SD)	63.1 (5.8)	57.1 (8.0)	1.14	1.12	1.14	<1x10 ⁻¹⁵
Mean age at diagnosis (SD)	67.7 (7.2)					
Townsend Deprivation Index (SD)	-0.78 (3.31)	-1.44 (3.05)	1.09	1.06	1.11	1.5x10 ⁻¹³
Pollution, NO ₂ µg/m ³ (SD)	27.1 (7.5)	26.3 (7.5)	1.02	1.01	1.03	5.1x10 ⁻⁵
Smoking status						
Never smoker (N, %)	182 (25.4%)	84,584 (48.5%)				
Former smoker (N, %)	412 (57.4%)	67,823 (38.9%)	2.10	1.76	2.50	<1x10 ⁻¹⁵
Current smoker (N, %)	108 (15.0%)	19,044 (10.9%)	2.83	2.23	3.60	<1x10 ⁻¹⁵
Missing (N, %)	16 (2.2%)	2,803 (1.2%)				
Body Mass Index (SD)	28.5 (4.7)	27.8 (4.2)	1.04	1.03	1.06	1.1x10 ⁻⁵
Median income band (IQR)	2 (1-3)	3 (2-4)	0.73	0.68	0.80	5.6x10 ⁻¹⁴
Mean FEV1 L (SD)	2.60 (0.67)	3.24 (0.75)	0.44	0.38	0.49	<1x10 ⁻¹⁵
FEV1 percent predicted (SD)	94.3 (22.1)	106.6 (20.7)	0.98	0.97	0.98	<1x10 ⁻¹⁵
Mean FVC L (SD)	3.53 (0.81)	4.31 (0.92)	0.47	0.42	0.52	<1x10 ⁻¹⁵
FVC percent predicted (SD)	107 (21.6)	120 (21.4)	0.97	0.97	0.98	<1x10 ⁻¹⁵
Metabolic Equivalent Task score (SD)	7.11 (1.32)	7.45 (1.16)	0.79	0.74	0.84	2.6x10 ⁻¹³
Adjusted T/S ratio (SD)	0.75 (0.13)	0.82 (0.13)	0.018	0.009	0.037	<1x10 ⁻¹⁵
Z-adjusted T/S log	-0.74 (1.19)	-0.11 (0.99)	0.63	0.58	0.67	<1x10 ⁻¹⁵
Participants deceased, N (%)	313 (43.6%)	7,838 (4.5%)	11.4	9.7	13.2	<1x10 ⁻¹⁵

Tables 2: Demographics for unrelated male idiopathic pulmonary fibrosis cases and controls of European ancestry in UK Biobank derived using logistic regression analyses. Odds ratios show the extent to which the exposures predicted prevalence of IPF, except for 'participants deceased', which shows the risk ratio for death with IPF. (Odds ratios and p values are adjusted for age, other than percent predicted spirometry which already includes adjustment). Income bands indicate average household income: (1) <£18000, (2) £18000-£30999, (3) £31000-51999, (4) £52000-£100000 and (5)>£100000.