Telomere Length and Risk of Idiopathic Pulmonary Fibrosis and Chronic Obstructive Pulmonary Disease: A Mendelian Randomisation Study

Anna Duckworth PhD [1, 2], Michael A. Gibbons PhD [1, 2, 3], Richard J. Allen PhD[4], Howard Almond [2], Robin N. Beaumont PhD [1], Andrew R. Wood PhD [1], Professor Katie Lunnon PhD [1], Professor Mark A. Lindsay PhD [5], Professor Louise V. Wain PhD [4, 6], Jess Tyrrell PhD [1, 7], Chris J. Scotton PhD [1, 2, 7]

[1] Institute of Biomedical & Clinical Science, College of Medicine & Health, University of Exeter, Exeter, UK

[2] Exeter Patients in Collaboration (EPIC) for PF, Exeter, UK

[3] Respiratory Medicine Department, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

[4] Department of Health Sciences, University of Leicester, Leicester, UK

[5] Department of Pharmacy and Pharmacology, University of Bath, Bath, UK

[6] National Institute for Health Research, Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, UK.

[7] These authors contributed equally

Name and address for correspondence:

Dr Chris Scotton

c.j.scotton@exeter.ac.uk

Institute of Biomedical & Clinical Science

College of Medicine & Health

University of Exeter

St Luke's Campus

Madgalen Road

Exeter

EX1 2LU

Tel: +44(0)1392722934

Keywords: idiopathic pulmonary fibrosis, chronic obstructive pulmonary disease, Mendelian randomisation, telomeres, genetics

Summary

Background

Idiopathic pulmonary fibrosis (IPF) is a fatal lung disease accounting for 1% of UK deaths. In the familial form of pulmonary fibrosis, causal genes have been identified in ~30% of cases, and a majority relate to telomere maintenance. Prematurely shortened leukocyte telomere length associates with IPF, and also chronic obstructive pulmonary disease (COPD), a disease with similar demographics and shared risk factors. Using Mendelian randomisation (MR), we investigated evidence supporting a causal role for short telomeres in IPF and COPD.

Methods

MR inference of telomere length causality was performed for IPF (up to 1,369 cases) and COPD (13,538 cases) against 435,866 controls of European ancestry in UK Biobank. Polygenic risk scores, followed by two-sample MR analyses were carried out using seven genetic variants previously associated with telomere length, with replication analysis in an IPF cohort (2,668 cases vs 8,591 controls) and COPD cohort (15,256 cases vs 47,936 controls).

Findings

In the UK Biobank, a genetically instrumented one standard deviation shorter telomere length was associated with higher odds of IPF (OR=4.19; 95% CI 2.33-7.55, P=0.0031) but not COPD (OR = 1.07; 95% CI 0.88-1.30, P = 0.51). Similarly, an association was found in the IPF replication cohort (OR = 12.3; 95% CI: 5.05-30.1, P = 0.015) and not in the COPD replication cohort (OR = 1.04; 95% CI 0.71-1.53, P = 0.83). Meta-analysis of the two-sample MR results provided evidence inferring that shorter telomeres cause IPF (5.81 higher odds of IPF; 95% CI 3.56-9.50; $P=2.19 \times 10^{-12}$). There was no evidence to infer that telomere length caused COPD (OR=1.07; 95% CI 0.90-1.27, P=0.46).

Interpretation

Cellular senescence is hypothesised as a major driving force in IPF and COPD; telomere shortening may be a contributory factor in IPF, suggesting divergent mechanisms in COPD. This enables greater focus in telomere-related diagnostics, treatments and the search for a cure in IPF. Therapies manifesting improvements in telomere length warrant investigation.

Funding

GW4 MRC DTP.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a complex and incurable fibrotic lung disease. Based on data from 2012 [1], the average lifespan following diagnosis is widely quoted as being around 3yrs and IPF accounts for around 5300 UK deaths each year, over 1% of all UK deaths.

Research in context

Evidence before this study

We searched in PubMed for publications up to 25th June 2020 on this subject using the search terms "telomere length", "Mendelian random*" and "pulmonary fibrosis" or "interstitial lung disease". We found a 2017 study which reports reduced risk of interstitial lung diseases associated with long telomeres (and no association for COPD) as part of a wider Mendelian Randomisation study on the association between telomere length and risk of cancer and non-neoplastic diseases. We found no further evidence for the use of MR to investigate causality in IPF. We searched in PubMed up to 10th June 2019 for GWAS results of SNPs associated with telomere length in order to create the instrument used in our MR analysis. We summarise the available evidence of association between telomere length and pulmonary fibrosis in animal models and in humans from publications up to this date.

Added value of this study

This study builds significantly on previous reports of a likely causal link between telomere length and IPF but not COPD by providing comprehensive evidence suggesting causality in IPF using large datasets from UK Biobank plus clinically well-defined replication cohorts; the study includes direct comparisons of underlying demographic features and draws on data from larger cohorts than have been used previously (1,133 IPFs in UKB; 2,668 IPFs in replication cohort; 11,413 COPDs in UKB; 15,256 COPDs in replication cohort).

Considerable experimental effort has been invested into finding a cause of PF in mouse models but there is no experimental evidence of causality in humans. This research therefore advances understanding in the fields of IPF and COPD research.

Implications of all the available evidence

Evidence inferring a cause for IPF allows research into more effective treatments to be pursued with renewed direction. It also allows clinicians to propose and patients to undertake known telomere maintenance therapies, which are also beneficial to general health.

The association of PF and COPD with telomere length

In approximately 20% of cases, pulmonary fibrosis (PF) clusters in family groups [2]. Inherited genetic causes have been established for around 30% of familial cases and the majority relate to telomere maintenance. The most common of these is the gene encoding telomerase reverse transcriptase (TERT), a catalytic subunit of the enzyme telomerase which works in conjunction with the telomerase RNA component, TERC [3]. Variants in these two genes account for 19% of familial cases via autosomal dominant inheritance with reduced penetrance.

Idiopathic pulmonary fibrosis (IPF) risk also has a strong genetic component [4]. Prematurely shortened leukocyte telomere length (LTL) has been associated with IPF and also chronic obstructive pulmonary disease (COPD), a condition with a similar

demographic. Studies have shown age adjusted LTL values of 0.85 ± 0.60 vs 1.15 ± 0.6, p=0.0001 relative to reference DNA for IPF versus controls [5] and 0.68 ± 0.25 vs. 0.88 ± 0.52, p = 0.003 for COPD versus smoking controls [6]. There is also evidence for an association between shorter LTL and worsened survival in IPF [7], with the suggestion that LTL could be used as a predictive biomarker. A recent study which included 32 IPF patients undergoing diagnostic lung biopsy demonstrated shortened lung telomeres (particularly in type II alveolar epithelial cells within fibrotic lesions), but no correlation with age; half of these showed excessive lung telomere shortening, to the same extent seen in PF patients with a TERT mutation – suggestive of a disease driven by telomere attrition [8].

The most widely used model for experimental exploration and pre-clinical assessment for IPF is the murine model of bleomycin-induced lung injury[9]. Although this mouse model has been an important precursor to clinical trials and has helped enable the development and licensing of the two main IPF treatments, Nintedanib and Pirfenidone, the majority of studies utilise young mice where the fibrotic lesions lack features which are characteristic of IPF (such as honeycombing and fibrotic foci), and where the fibrosis may resolve depending on the severity of initial injury [10]. Mice have particularly long telomeres; studies which have utilised aged mice (>18mo old) have demonstrated a more profound fibrotic response which does not resolve [11], and which better reflects the human disease. Notably, a recent study of low-dose bleomycin in *TERT*$^{-/-}$ mice (which have significantly shortened telomeres) demonstrated that targeted telomerase activation in type II alveolar epithelial cells (AEC2s) cells using gene therapy with adeno-associated vectors (AAV) showed therapeutic effects in mice with established fibrosis [12] through telomere elongation and increased proliferation of AEC2 cells combined with lower DNA damage, apoptosis and senescence burden.

Thus, while there is evidence of an association with shorter telomeres in both diseases, experimentally, there is currently only evidence from murine models of PF that telomere extension is therapeutic.

Mendelian Randomisation

A genetic technique known as Mendelian randomisation (MR) [13] can be used to test for inference of a causal relationship between a phenotype that can be genetically influenced (such as telomere length) and a disease outcome, such as IPF or COPD. Causality in one direction is inferred because genetic make-up is allocated at conception and unlikely to be influenced by disease in later life. Potential confounding influences such as smoking, pollution and other environmental/lifestyle risk factors are removed from the analysis, creating in effect a natural blind randomised control trial (Figure 1). Recent genome-wide association studies (GWAS) [14,15] have identified several genetic variants or single nucleotide polymorphisms (SNPs) that are independently associated with telomere length and provide potential tools for MR. A previous study investigating associations between genetically increased telomere length and risk of cancer and other non-neoplastic diseases, reported increased risk of site-specific cancers and reduced risk of coronary artery disease, celiac disease and interstitial lung diseases [16]. We therefore hypothesised that telomere length is causally linked to IPF but not COPD, given that inherited genetic defects in telomerase production lead to familial pulmonary fibrosis. To test this, we used MR and the latest data release from up to 451,025 participants in the UK Biobank together with a genetic instrument associated with shorter telomere length. We further tested the effect in males and females separately since there is a well-established gender bias in IPF.

Methods

Collection and selection of UK Biobank Data

The UK Biobank is a study of 500,000 volunteers aged between 37-73 years, recruited across the UK during 2006-2010 [17], with participant data including physical measurements, biological samples (blood, urine and saliva) for biomarker and genetic analysis, and long-term follow-up via hospital record linkage. Information on patient and public involvement is available online [18]. No further ethical or IRB approval was required. The cut-off date for data in this study was 31/3/2017. Genetic variant or single nucleotide polymorphism (SNP) data was generated from the Affymetrix Axiom UK Biobank array (for ~450,000 individuals) and the UK BiLEVE array (~50,000 individuals) following extensive quality control, as described previously [19]. Individuals were defined as European descent using principal component analysis (PCA). Briefly, principal components were generated using loadings from high-confidence SNPs in the 1000 Genomes Cohort. The loadings were then used to project all of the UK Biobank samples into the same principal component space, and individuals were then clustered using the first four principal components. We identified two UK Biobank study populations: 1) The full set of available participants (451,025 individuals of European ancestry), including related individuals, to maximise statistical power in the 2-sample Mendelian randomisation where relatedness is handled within the model; 2) A smaller subset of 379,708 unrelated individuals (defined using a KING Kinship, generating an optimal list of unrelated individuals with maximum inclusion), important for the initial regression modelling. Ancestral principal components were then generated within these identified individuals for use in subsequent analyses, as described previously [20].

IPF cases within UK Biobank were defined as those having a primary or secondary ICD10 code HES (Hospital Episodes Statistics) diagnosis of J84.1. With this 'narrow' criterion, we identified 1369 cases (1133 unrelated, whereby no two participants are third degree related or closer). We repeated our analysis with a 'broad' IPF definition that included J84.0 (Alveolar and parieto-alveolar conditions), J84.8 (Other specified interstitial pulmonary diseases) or J84.9 (Interstitial pulmonary disease, unspecified) and using these criteria, we identified 1,621 (1,353 unrelated) cases. COPD cases were defined as those having a primary or secondary ICD10 code of J41 (Simple and mucopurulent chronic bronchitis), J42 (Unspecified chronic bronchitis), J43 (Emphysema) or J44 (Other chronic obstructive pulmonary disease) plus those self-reported to have COPD: a total of 14,103 (11,895 unrelated) cases. Using the broad definition of IPF, we subtracted 565 (482 unrelated) cases with both IPF and COPD to leave 13,538 (11,413 unrelated) cases.

For the control group, we removed both COPD cases and broad definition IPF cases to give 435,866 (366,942 unrelated) clean controls.

Replication cohorts

An IPF replication cohort was derived from the discovery stage of a recent GWAS study which comprised three independent case-control studies from the UK, Chicago and Colorado a total of 2,668 IPF cases and 8,591 controls [21]. The UK study included matched controls selected from UK Biobank.

The COPD replication cohort was derived from a recent COPD GWAS study with 15,256 cases and 47,936 controls, with 13,710/15,256 (90%) of cases and 38,062/47,936 (79%) of controls being of European ancestry [22].

IPF and COPD diagnoses were made in accordance with accepted international criteria. Both replication studies had appropriate international review board or ethics approval.

Observational Associations

We used logistic regression models in Stata 13.0 to compare the key demographics of the IPF and COPD groups with controls, adjusting for age and sex as appropriate. These analyses were performed in the unrelated subset of individuals to prevent familial bias.

Identification of genetic instrument variants

Genetic variants for telomere length were chosen from published GWAS studies, using GWAS that did not include data from the UK Biobank. We used an instrument composed of 7 variants robustly associated with LTL derived from a genome-wide meta-analysis of 37,684 individuals of European descent, with replication of selected variants in a further 10,739 individuals by Codd *et al* [14] (see Supplement p1). Other telomere length GWAS results were available (see [16]) but with smaller sample sizes and several of the variants available were in linkage disequilibrium (Supplement p1) so did not meet our MR criteria.

The extracted genetic variants, recoded to 0, 1 or 2 according to the number of telomere length associated alleles, were used to create a genetic risk score (GRS) for telomere length for each individual. The variants were weighted by their effect size (ß-coefficient) obtained from the primary GWAS, where each beta value reflects the standard deviation change (1-SD) in LTL per copy of the effect allele:

$$Weighted\ score = ß_1 \times SNP_1 + ß_2 \times SNP_2 + \ldots\ldots ß_n \times SNP_n$$

Mendelian Randomisation

Mendelian randomisation (MR) was used to investigate causality between telomere length and incidence of IPF and COPD. MR relies on several general assumptions [23] which are applied in this case as follows:

a) the telomere length genetic variants are robustly associated with absolute LTL
b) the telomere length genetic variants are not associated, independently of their effects on telomere length, with confounding variables for IPF or COPD
c) the telomere length genetic variants are only associated with IPF or COPD via their effect on telomere length

In this study, we employed several MR approaches. First, we investigated the association between IPF or COPD and the telomere length genetic risk scores in the unrelated data set of 379,708 individuals using logistic regression models. Ancestral principal components (as previously described [24]) were included as covariates in the analysis to control for residual population structure and we also adjusted for baseline age, sex and UK Biobank assessment centre.

We next performed two-sample MR using the BOLT-LMM (v2.3.4) algorithm for mixed model association testing. The seven genetic variants associated with telomere length were extracted. A standard Inverse Variance Weighted (IVW) instrumental variable analysis was performed in R (v3.5.2), along with two methods that are more resistant to pleiotropy: MR-Egger [25] and Median MR [26]. The IVW method regresses the effect sizes of variant outcome associations (here telomere length associated variants vs incidence of IPF or COPD) against effect sizes of the variant risk factor associations (here telomere length associated variants vs telomere length). Variant risk factor associations were taken from the primary GWAS of telomere length [14]. If no heterogeneity is detected amongst the causal estimates, the IVW analysis is carried out under a fixed effect model with the assumption of no horizontal pleiotropy. Alternatively, if heterogeneity is found amongst the causal estimates, a random effects model is implemented and the approach assumes that:

a) Either the strength of the association of the genetic instruments with the risk factor is not correlated with the magnitude of the pleiotropic effects
b) Or the pleiotropic effects have an average value of zero

In contrast, the MR-Egger method uses a weighted regression with an unconstrained intercept to remove the assumption that all genetic variants are valid instrumental variables and is therefore less susceptible to confounding from potentially pleiotropic variants that have a stronger effect on the incidence of IPF compared to their effect on telomere length. The Median-MR method is also more resistant to pleiotropy; it takes the median instrumental variable from all variants included and is therefore robust when <50% of the genetic variants are invalid. Given these different assumptions, if all methods are broadly consistent this strengthens the causal inference. Details of the R code for the 2-sample IVW, MR-Egger and Median-MR analyses are available in Bowden *et al* [26,27].

Two sample MR in our IPF and COPD replication cohorts used summary GWAS statistics, the majority of which excluded UK Biobank data (7,893/11,259=70% for IPF and 100% FOR COPD).

Additional sensitivity analyses were performed to check the robustness of our results. First, the MR was repeated, excluding rs2736100 which has a reported link with IPF [28]. Second, we repeated the regression analyses in the unrelated subset of individuals using sex specific genetic risk scores (see Supplement p2). Third, we checked sensitivity of regression analyses for COPD to case description, since the definition of COPD cases in our replication cohort was based on lung function data (FVC% predicted <80%, FEV1/FVC <0.7) rather than reported diagnosis and recent evidence has shown that a spirometric definition identifies more cases [29]. Percent predicted spirometry values were generated in UK Biobank data using the Global Lung Initiative reference equations, for which the methodology used to derive the lower limit of normal (LLN) for all equations takes into account that the spread of values around the predicted values is not normal and depends on age and outcome [30].

Role of the funding source

The funding source had no involvement in: study design; collection, analysis, and interpretation of data; writing of the report; the decision to submit the paper for publication. AD, RNB, ARW, 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 demographics of the unrelated 1,133 IPF 'narrow' cases and 11,413 COPD cases are summarised in Table 1 and 2. In general the odds ratios quoted for each exposure predict prevalence of IPF or COPD, but the mortality figures are the risk ratio for death with each disease. Strong associations were noted between IPF and a range of demographic and environmental variables. Briefly, older age, male sex, lower socioeconomic position, ever smoking, reduced lung function and reduced exercise were associated with a higher odds ratio for IPF and also for COPD. Similar associations were noted for the 1,353 IPF 'broad' cases and also in the larger cohort of related individuals (Supplement p3).

Telomere length polygenic genetic risk score

Using the 7-variant telomere length genetic risk score in our UK Biobank unrelated cohort, we demonstrated that the telomere length GRS was associated with higher odds of IPF in our 'narrow' definition group of 1,133 IPF cases; odds ratio (OR) = 1.11 [95%CI: 1.07-1.15], P = 2.1×10^{-8} (Figure 2). Repeating this for the broader definition IPF, containing 1,353 cases, did not greatly alter this finding, OR = 1.09 [95%CI: 1.05-1.13], P = 4.7×10^{-7} (Table 3).

In contrast, no association was seen for our COPD group of 11,413 cases (OR 1.000 [95%CI: 0.99-1.01], P=0.98; Figure 2).

With the 6 variant GRS (excluding rs2736100), the results were slightly attenuated but the GRS remained associated with IPF; for the 'narrow' group OR = 1.08 [95%CI: 1.04-1.13], P = 2.5×10^{-4} and for the broader definition IPF group OR = 1.06 [95%CI: 1.02-1.10], P = 0.0027.

Due to the strong association between IPF incidence and male sex (OR = 1.94 [95%CI: 1.71-2.19], P <1×10^{-15}; Table 1), we used separate genetic risk scores for each sex, created from the 7 SNP sex specific beta values for telomere length. Repeating the MR generated similar results for males and females (Table 3).

In our sensitivity studies, we repeated the analysis for COPD using a lung function definition and the results were very similar (Table 3).

Two-sample MR analysis

Two sample MR provided evidence that genetic predictors of telomere length are ssociated with the risk of IPF. From this, under untestable assumptions, we infer that shorter telomere length is a cause of IPF. In the UK Biobank, a genetically instrumented one standard deviation shorter telomere length was associated with higher odds of IPF; using the IVW method, in the narrow IPF group, OR = 4.19 [95% CI 2.33-7.55], P = 0.0031 and in the broad IPF group OR = 3.28 [95% CI 1.77-6.08], P = 0.0093. There was no evidence of a causal relationship in COPD (OR = 1.07 [95% CI 0.88-1.30], P = 0.51; Figures 3 & 4 and Supplement p5). For the IPF groups, all methods were directionally consistent and the Egger method provided no evidence of pleiotropy (Egger intercept P-value: 0.45 in IPF 'narrow' and 0.47 in IPF 'broad'). There was no evidence of SNP heterogeneity for either cohort (Supplement p5)

MR results from replication cohorts

To provide further evidence for or against the causal role of telomere length in IPF, we used the summary statistics from the replication cohort data from 2,668 IPF patients and 8,591 controls. Again, two-sample MR provided evidence inferring a causal

role for shortened telomeres in IPF. A 1-SD decrease in telomere length was associated with OR = 12.3 [95% CI: 5.05-30.1], P = 0.0015 (Figures 3 & 4 & Supplement p5). Similar associations were noted using the MR methods that are more robust to pleiotropy, and the Egger intercept provided no evidence of pleiotropy (P-value: 0.75).

Results from the COPD replication cohort provided no evidence of a causal relationship in COPD (OR = 1.04 [95% CI 0.71-1.53], P = 0.83; Figures 3 & 4 and Supplement p5).

There was some evidence of heterogeneity of the SNP effects (Supplement p5), with rs7675998 driving this heterogeneity in the IPF cohort. When the analysis was repeated without this SNP, the finding that short telomeres predict IPF remained.

Meta-analysis of UK Biobank and replication cohort data

Meta-analysis of the IVW model estimates for IPF from the UK Biobank and the replication cohort also inferred a causal relationship with a 1-SD shorter telomere length leading to OR= 5.81 [95% CI: 3.56-9.50], P= 2.19×10^{-12} for IPF. Using this model there is some evidence of heterogeneity between cohorts (heterogeneity P= 0.05). In comparison, meta-analysis of IVW estimates for COPD from the UK Biobank and the COPD replication cohort provided no evidence of any causal relationship for telomere length, with OR= 1.07 [95%CI:0.90-1.27, P=0.46, heterogeneity P=0.91]. Independent cohort and combined meta-analysis results are shown in Figure 4.

Similarly, meta-analysis results for MR-Egger estimates for IPF inferred a causal relationship for telomeres in IPF with a 1-SD shorter telomere length leading to OR= 15.8 [95% CI: 1.48-169], P= 0.022, with heterogeneity P= 0.80. Again, meta-analysis results using MR-Egger estimates for COPD provided no evidence of any causal relationship for telomere length, with OR= 1.07 [95%CI:0.45-2.56, P=0.88, heterogeneity P=0.87]. Independent cohort and combined meta-analysis results are shown in Supplement p8.

Discussion

Using a Mendelian randomisation approach, we have shown that decreased telomere length is associated with increased risk of IPF but not COPD. The majority of our findings persisted when they were derived using models that made allowance for violations of MR assumptions, such as confounding by pleiotropy, with outcomes which were broadly consistent. An earlier study indicated that some genetic variants regulating telomere length are associated with risk of IPF and not COPD [16] and we further build on that assertion with data from larger cohorts than have been used previously, with a comprehensive and focused study, where we see a dose-dependent relationship in IPF across seven genetic predictors of telomere length. We also see little evidence of heterogeneity between variants in our cohorts. In the one instance where heterogeneity in the IPF replication cohort was apparent, this was driven by the single outlier SNP (rs7675998), which appears to have a different association with disease in UK and US cohorts. Taken together, these data therefore provide robust evidence inferring a causal link from short telomeres for idiopathic pulmonary fibrosis, while suggesting divergent underlying disease mechanisms in COPD. While COPD is not a reported association in families with telomere biology disorders [31], there are indications that short telomeres lower the threshold of chronic cigarette smoke-induced damage [32]. Thus, although short telomeres do not appear to have a causal role in COPD, they may contribute to the age-related onset of emphysema.

Both IPF and COPD are exemplars of age-related disease, with increasing focus on the impact of accelerated ageing and particularly the pathogenic role of cellular senescence (see [33] for a comprehensive review). This is a complex cellular programme and can be broadly subdivided into replicative and cellular senescence [34]; the former results from intrinsic cellular events, including telomere shortening, while the latter can be driven by various stimuli including oxidative stress and DNA damage [35]. IPF and COPD both have significantly shorter LTL [5,6]; shortened telomeres have also been identified directly within lung tissue from IPF patients [8,36] but not in COPD [37]. While there have been mixed reports on the direct correlation between leukocyte and tissue telomere length generally [37-39], the study by van Batenburg *et al* reported a significant positive correlation between LTL and lung biopsy telomere length in IPF (n=32, r=0.53, p=0.002) [8]. In light of our findings, we speculate that telomere shortening could act as an intrinsic and systemic driving force of cellular senescence in IPF, akin to the relationship between telomere-associated driver mutations in familial PF. In both diseases, short LTL could also result from immunosenescence, or possibly hypoxemia, given the reported association between telomere length and PaO_2 [40]. Notably, LTL in COPD does not appear to be influenced by smoking status (current vs ex-smokers [41]), although smoking *per se* can cause telomere attrition [42]. Of interest, a recent study by Kachuri *et al* [43] used MR to investigate a causal role for longer telomeres in lung cancer, demonstrating an increased odds ratio for lung adenocarcinoma but not squamous cell carcinoma (SCC) or head and neck cancer. IPF is linked with an increased risk of lung cancer [44], but notably SCC rather than adenocarcinoma [45].

Evidence inferring a causal role for premature telomere attrition in IPF does present several significant opportunities. Approaches aimed at addressing inadequate telomere maintenance either universally or in the cells contributing most to the disease pathogenesis, may offer a therapeutic option. Restoration of telomere length is not as straightforward as simply upregulating telomerase since this would promote the risk of cancers, although clinical trials are underway using this approach in other conditions (see NCT04110964 at clinicaltrials.gov). In light of this, safe telomerase activation therapy is being explored in other medical fields such as cardiology, using transient delivery which avoids creating an environment in which increased telomerase persists [46]. Targeted delivery *in vivo* presents a very significant challenge to utilising such an approach in PF, but future developments may render this feasible. Diminution of the senescent cell burden in IPF is also currently being explored using senolytic drugs such as the combination of dasatinib plus quercetin, for which an open label Phase I Study has recently been reported [47]. Alternative approaches could include exploration of androgen therapy: testosterone has been used effectively to treat the telomere linked disease, aplastic anaemia, for many years and Danazol has shown some promise in the treatment of PF, with an apparent arrest in lung function decline over 36 months [48]. Androgens can also restore telomerase to normal levels in cells from telomere disease patients who are heterozygous for TERT gene mutations [49].

There are also approaches which improve general wellbeing while also having a positive effect on promoting telomere length. We speculate these could be further promoted as an important part of clinical management of IPF patients, and include exercise [50], reduction of life stress [51] and mindfulness [52,53]. Our data show that participants with IPF tended to have reduced physical activity (Table 1), probably exaggerated by poor physical health; carefully increased exercise could offer multiple benefits in terms of promoting chromosomal telomere length and also boosting fitness and mental health. Similarly, stress arising from difficult life circumstances such as social deprivation (which our data show is associated with IPF, Table 1) can be relieved with

practices such as mindfulness [52,53]. Relatively simply practices which improve patient wellbeing may also have a fundamentally positive impact on reducing telomere attrition.

The most established genetic association with IPF to date has been a polymorphism (rs35705950) in the promoter of the MUC5B gene, and an inverse relationship between the presence of the risk allele and the presence of telomerase pathway rare variants has been reported [4]. Although we had insufficient patient numbers to investigate this in the current study, stratification by MUC5B SNP status would be worthy of exploration in future work.

Our study does have several limitations. While numbers of cases and controls in UK Biobank are large, those cases lack extensive clinical characterisation. This is mitigated to some extent by the thorough adjudicated diagnoses for cases in both our replication cohorts. Second, we have not assessed LTL directly since those data were not available for these cohorts. Instead, we have assumed that any systemic modifiers of telomere length will affect cells in which telomere length is actively maintained throughout life. This will include haematopoietic stem cells in the bone marrow from which leukocytes derive, but also type II alveolar epithelial cells, which are a focus of damage in the IPF lung. Our study does not therefore attempt to address the absolute extent to which telomere length in IPF is genetically or environmentally determined. Mendelian randomisation also has inherent limitations, in light of the three key assumptions described in the Methods. We have tested these assumptions using the different models described herein and, while the results were found to be broadly consistent, some residual uncertainty inevitably remains. This can be reduced further in future analyses with larger sample size, as and when these become available.

To conclude, we have found evidence, in both our UK Biobank and replication cohorts, to infer a causal link between telomere length and sporadic IPF. We found no evidence of any link to COPD, a similar age-related disease, in either the UK Biobank or our COPD replication cohort. The inference of a cause behind human IPF leads us to new insights towards beneficial therapies for patients and routes to potential new treatments - which may lead us closer to preventing the disease in those with prematurely shortened telomeres, and ultimately providing a direction in our search for a cure.

Author Contributions

AD conceived the project, carried out the analyses, interpreted the results and drafted the manuscript; MAG provided disease cohort definitions and clinical guidance throughout; HA provided comment and feedback from the patient perspective; RNB and ARW defined and carried out supporting analyses; KL and MAL provided supervisory support to AD; RJA and LVW provided the replication cohort summary data; 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

AD is funded by the GW4 BioMed MRC Doctoral Training Partnership. 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]. RJA is an Action for Pulmonary Fibrosis Research Fellow. LVW holds a GSK/British Lung Foundation Chair in Respiratory Research. This research has been conducted using the UK Biobank Resource (applications 9072 and 44046). The authors would like to acknowledge the use of the University of Exeter High-Performance Computing (HPC) facility in carrying out this work. The Research was partially supported by the National Institute for Health Research (NIHR) Leicester Biomedical Research Centre; the views expressed are those of the author(s) and not necessarily those of the National Health Service (NHS), the NIHR or the Department of Health.

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, RJA, HA, RNB, ARW, KL, MAL, JT and CJS have no declaration of interest.

References

1. British Lung Foundation. The battle for breath: the impact of lung disease in the UK, 2016.
2. García-Sancho C, Buendía-Roldán I, Fernández-Plata MR, et al. Familial pulmonary fibrosis is the strongest risk factor for idiopathic pulmonary fibrosis. *Respiratory Medicine* 2011; **105**(12): 1902-7.
3. 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.
4. Dressen A, Abbas AR, Cabanski C, et al. Analysis of protein-altering variants in telomerase genes and their association with MUC5B common variant status in patients with idiopathic pulmonary fibrosis: a candidate gene sequencing study. *The Lancet Respiratory Medicine* 2018; **6**(8): 603-14.
5. 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.
6. Córdoba-Lanús E, Cazorla-Rivero S, Espinoza-Jiménez A, et al. Telomere shortening and accelerated aging in COPD: findings from the BODE cohort. *Respiratory research* 2017; **18**(1): 59.
7. 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.
8. van Batenburg AA, Kazemier KM, van Oosterhout MFM, et al. From organ to cell: Multi-level telomere length assessment in patients with idiopathic pulmonary fibrosis. *PLOS ONE* 2020; **15**(1): e0226785.
9. Tashiro J, Rubio GA, Limper AH, et al. Exploring Animal Models That Resemble Idiopathic Pulmonary Fibrosis. *Frontiers in Medicine* 2017; **4**(118).
10. Scotton CJ, Hayes B, Alexander R, et al. *Ex vivo* micro-computed tomography analysis of bleomycin-induced lung fibrosis for preclinical drug evaluation. *European Respiratory Journal* 2013; **42**(6): 1633-45.
11. Hecker L, Logsdon NJ, Kurundkar D, et al. Reversal of persistent fibrosis in aging by targeting Nox4-Nrf2 redox imbalance. *Sci Transl Med* 2014; **6**(231): 231ra47-ra47.
12. Povedano JM, Martinez P, Serrano R, et al. Therapeutic effects of telomerase in mice with pulmonary fibrosis induced by damage to the lungs and short telomeres. *eLife* 2018; **7**: e31299.
13. Smith GD. Mendelian Randomization for Strengthening Causal Inference in Observational Studies:Application to Gene × Environment Interactions. *Perspectives on Psychological Science* 2010; **5**(5): 527-45.
14. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. *Nature genetics* 2013; **45**(4): 422-7e4272.
15. Mangino M, Hwang S-J, Spector TD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. *Hum Mol Genet* 2012; **21**(24): 5385-94.

16. The Telomeres Mendelian Randomization Collaboration. Association Between Telomere Length and Risk of Cancer and Non-Neoplastic Diseases: A Mendelian Randomization Study. *JAMA oncology* 2017; **3**(5): 636-51.
17. 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.
18. UK Biobank Ethics and Governance Framework. Available from: https://www.ukbiobank.ac.uk/the-ethics-and-governance-council/ (accessed 02/07/2020)
19. UK Biobank. https://www.ukbiobank.ac.uk/scientists-3/genetic-data/ (accessed 02/07/2020)
20. Tyrrell J, Mulugeta A, Wood AR et al. Using genetics to understand the causal influence of higher BMI on depression. *Int J Epidemiol* 2019; **48**(3):834-848.
21. Allen RJ, Guillen-Guio B, Oldham JM, et al. Genome-Wide Association Study of Susceptibility to Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine* 2020; **201**(5):564-574.
22. Hobbs BD, de Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nature Genetics* 2017; **49**: 426.
23. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine* 2008; **27**(8): 1133-63.
24. UK Biobank Genetic Data: MRC-IEU Quality Control, version 3. 2018. https://data.bris.ac.uk/data/dataset/1ovaau5sxunp2cv8rcy88688v (accessed 02/07/2020)
25. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *International journal of epidemiology* 2016; **45**(6): 1961-74.
26. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic epidemiology* 2016; **40**(4): 304-14.
27. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology* 2015; **44**(2): 512-25.
28. Mushiroda T, Wattanapokayakit S, Takahashi A, et al. A genome-wide association study identifies an association of a common variant in *TERT* with susceptibility to idiopathic pulmonary fibrosis. *Journal of Medical Genetics* 2008; **45**(10): 654-6.
29. Sakornsakolpat P, Prokopenko D, Lamontagne M, et al. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nature genetics* 2019; **51**(3): 494-505.
30. Cooper BG, Stocks J, Hall GL, et al. The Global Lung Function Initiative (GLI) Network: bringing the world's respiratory reference values together. *Breathe (Sheff)* 2017; **13**(3): e56-e64.

31. Savage SA, Bertuch AA. The genetics and clinical manifestations of telomere biology disorders. *Genet Med* 2010; **12**(12): 753-64.
32. Alder JK, Guo N, Kembou F, et al. Telomere length is a determinant of emphysema susceptibility. *American journal of respiratory and critical care medicine* 2011; **184**(8): 904-12.
33. Cho SJ, Stout-Delgado HW. Aging and Lung Disease. *Annual Review of Physiology* 2020; **82**(1): null.
34. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature* 1990; **345**(6274): 458-60.
35. Robles SJ, Adami GR. Agents that cause DNA double strand breaks lead to p16INK4a enrichment and the premature senescence of normal fibroblasts. *Oncogene* 1998; **16**(9): 1113-23.
36. Snetselaar R, van Batenburg AA, van Oosterhout MFM, et al. Short telomere length in IPF lung associates with fibrotic lesions and predicts survival. *PloS one* 2017; **12**(12): e0189467.
37. Everaerts S, Lammertyn EJ, Martens DS, et al. The aging lung: tissue telomere shortening in health and disease. *Respiratory Research* 2018; **19**(1): 95.
38. Dlouha D, Maluskova J, Kralova Lesna I, Lanska V, Hubacek JA. Comparison of the relative telomere length measured in leukocytes and eleven different human tissues. *Physiological research* 2014; **63 Suppl 3**: S343-50.
39. Friedrich U, Griese E-U, Schwab M, Fritz P, Thon K-P, Klotz U. Telomere length in different tissues of elderly patients. *Mechanisms of Ageing and Development* 2000; **119**(3): 89-99.
40. Savale L, Chaouat A, Bastuji-Garin S, et al. Shortened Telomeres in Circulating Leukocytes of Patients with Chronic Obstructive Pulmonary Disease. *American Journal of Respiratory and Critical Care Medicine* 2009; **179**(7): 566-71.
41. Lee J, Sandford AJ, Connett JE, et al. The Relationship between Telomere Length and Mortality in Chronic Obstructive Pulmonary Disease (COPD). *PLOS ONE* 2012; **7**(4): e35567.
42. Morlá M, Busquets X, Pons J, Sauleda J, MacNee W, Agustí AGN. Telomere shortening in smokers with and without COPD. *European Respiratory Journal* 2006; **27**(3): 525-8.
43. Kachuri L, Saarela O, Bojesen SE, et al. Mendelian Randomization and mediation analysis of leukocyte telomere length and risk of lung and head and neck cancers. *International Journal of Epidemiology* 2018; **48**(3): 751-66.
44. HUBBARD R, VENN A, LEWIS S, BRITTON J. Lung Cancer and Cryptogenic Fibrosing Alveolitis. *American Journal of Respiratory and Critical Care Medicine* 2000; **161**(1): 5-8.
45. Park J, Kim DS, Shim TS, et al. Lung cancer in patients with idiopathic pulmonary fibrosis. *European Respiratory Journal* 2001; **17**(6): 1216-9.
46. Ramunas J, Yakubov E, Brady JJ, et al. Transient delivery of modified mRNA encoding TERT rapidly extends telomeres in human cells. *The FASEB Journal* 2015; **29**(5): 1930-9.

47. Justice JN, Nambiar AM, Tchkonia T, et al. Senolytics in idiopathic pulmonary fibrosis: Results from a first-in-human, open-label, pilot study. *EBioMedicine* 2019; **40**: 554-63.

48. Townsley DM, Dumitriu B, Liu D, et al. Danazol Treatment for Telomere Diseases. *The New England journal of medicine* 2016; **374**(20): 1922-31.

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. Cherkas LF, Hunkin JL, Kato BS, et al. The association between physical activity in leisure time and leukocyte telomere length. *Archives of Internal Medicine* 2008; **168**(2): 154-8.

51. Epel ES, Blackburn EH, Lin J, et al. Accelerated telomere shortening in response to life stress. *Proceedings of the National Academy of Sciences of the United States of America* 2004; **101**(49): 17312-5.

52. Jacobs TL, Epel ES, Lin J, et al. Intensive meditation training, immune cell telomerase activity, and psychological mediators. *Psychoneuroendocrinology* 2011; **36**(5): 664-81.

53. Schutte NS, Malouff JM. A meta-analytic review of the effects of mindfulness meditation on telomerase activity. *Psychoneuroendocrinology* 2014; **42**: 45-8.

Tables

Demographic	IPF J84.1 only	Controls	OR	95% CI		P
N = 368,075	1,133	366,942				
Mean age at baseline (SD) Mean age at diagnosis (SD)	63.1 (5.8) 67.2 (7.5)	57.1 (8.0)	1.13	1.12	1.14	$<1X10^{-15}$
Male sex, N (%)	718 (63.4%)	167,910 (45.8%)	1.94	1.71	2.19	$<1X10^{-15}$
Townsend Deprivation Index (SD)	-0.80 (3.33)	-1.53 (2.95)	1.10	1.08	1.12	$<1X10^{-15}$
Pollution, NO_2 $\mu g/m^3$ (SD)	27.2 (8.0)	26.2 (7.4)	1.02	1.02	1.03	$6.0x10^{-10}$
Smoking status						
Never smoker	341 (30.1%)	201,987 (55.1%)				
Former smoker	602 (53.1%)	128,242 (35.0%)	2.14	1.87	2.45	$<1X10^{-15}$
Current smoker	167 (14.7%)	32,022 (8.73%)	3.42	2.83	4.13	$<1X10^{-15}$
Missing	23 (2.0%)	5,104 (1.3%)				
Median income band (IQR)	2 (1-3)	3 (2-4)	0.69	0.65	0.74	$<1X10^{-15}$
Mean FEV1 L (SD)	2.36 (0.69)	2.77 (0.77)	0.35	0.32	0.40	$<1X10^{-15}$
FEV1 percent predicted (SD)	85.0 (23.5)	91.1 (22.6)	0.99	0.98	0.99	$<1X10^{-15}$
Mean FVC L (SD)	3.18 (0.88)	3.66 (1.00)	0.42	0.38	0.46	$<1X10^{-15}$
FVC percent predicted (SD)	114.8 (29.6)	120.2 (29.6)	0.99	0.99	1.00	$<1X10^{-15}$
Physical activity score (SD)	7.07 (1.31)	7.41 (1.13)	0.76	0.73	0.81	$<1X10^{-15}$
Participants deceased, N (%)	*439 (38.8%)*	*10,977 (3.0%)*	*11.8*	*10.4*	*13.4*	*$<1X10^{-15}$*

Table 1: Demographics for idiopathic pulmonary fibrosis cases using narrow J84.1 ICD10 code definition, from unrelated individuals 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 and sex, 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	COPD	Controls	OR	95% CI		P
N = 378,355	11,413	366,942				
Mean age at baseline (SD) Mean age at diagnosis (SD)	61.9 (6.2) 65.3 (7.4)	57.1 (8.0)	1.10	1.09	1.10	$<1X10^{-15}$
Male sex, N (%)	6,245 (54.7%)	167,910 (45.8%)	1.38	1.32	1.43	$<1X10^{-15}$
Townsend Deprivation Index (SD)	0.21 (3.6)	-1.53 (2.95)	1.19	1.19	1.20	$<1X10^{-15}$
Pollution NO_2 $\mu g/m^3$ (SD)	28.0 (7.7)	26.2 (7.4)	1.04	1.03	1.04	$<1X10^{-15}$
Smoking status						
Never smoker	1,806 (15.8%)	201,987 (55.1%)				
Former smoker	5,480 (48.0%)	128,242 (35.0%)	3.96	3.75	4.18	$<1X10^{-15}$
Current smoker	3,719 (32.6%)	32,022 (8.73%)	15.0	14.1	15.9	$<1X10^{-15}$
Missing	408 (3.6%)	5,104 (1.3%)				
Median income band (IQR)	1 (1-2)	3 (2-4)	0.53	0.52	0.54	$<1X10^{-15}$
Mean FEV L (SD)	2.01 (0.72)	2.77 (0.77)	0.15	0.14	0.15	$<1X10^{-15}$
FEV1 percent predicted (SD)	71.2 (23.7)	91.1 (22.6)	0.96	0.96	0.96	$<1X10^{-15}$
Mean FVC L (SD)	3.04 (0.93)	3.66 (1.00)	0.34	0.33	0.35	$<1X10^{-15}$
FVC percent predicted (SD)	107.6 (30.4)	120.2 (29.6)	0.98	0.98	0.98	$<1X10^{-15}$
Physical activity score (SD)	7.12 (1.31)	7.41 (1.13)	0.80	0.79	0.81	$<1X10^{-15}$
Participants deceased, N (%)	1,877 (16.5%)	*10,977 (3.0%)*	*4.59*	*4.35*	*4.85*	$<1X10^{-15}$

Table 2: Demographics for chronic obstructive pulmonary disease cases, from unrelated individuals of European ancestry in UK Biobank derived using logistic regression analyses Odds ratios show the extent to which the exposures predicted prevalence of COPD, except for 'participants deceased', which shows the risk ratio for death with COPD. (Odds ratios and p values are adjusted for age and sex, as above). Income bands indicate average household income: (1)<£18000, (2)£18000-£30999, (3)£31000-51999, (4) £52000-£100000 and (5)>£100000.

.

Case Group	Genetic instrument	N cases (controls)	Odds ratios (95%CI) for IPF vs control per SD change in LTL	P
IPF 'narrow' All	7 SNP	1,133 (366,942)	1.11 (1.07,1.15)	$2.1x10^{-8}$
IPF 'narrow' All	6 SNP - no rs2736100	1,133 (366,942)	1.08 (1.04, 1.13)	$2.5x10^{-4}$
IPF 'broad' All	7 SNP	1,353 (366,942)	1.09 (1.05, 1.13)	$4.9x10^{-7}$
IPF 'broad' All	6 SNP - no rs2736100	1,353 (366,942)	1.06 (1.02, 1.10)	$2.7x10^{-3}$
IPF 'narrow' Men only	7 SNP male grs	718 (174,254)	1.10 (1.05, 1.15)	$2.4x10^{-5}$
IPF 'narrow' Women only	7 SNP female grs	415 (204,321)	1.12 (1.05, 1.19)	$2.2x10^{-4}$
COPD (our definition)	7 SNP	11,413 (366,942)	1.000 (0.988, 1.012)	0.98
COPD (our definition)	6 SNP	11,413 (366,942)	1.002 (0.989, 1.016)	0.71
COPD (lung function definition)	7 SNP	30,359 (347,996)	1.003 (0.995, 1.010)	0.48

Table 3: Evidence from UK Biobank data suggesting a causal role for telomere length in IPF but not in COPD. Associations between shorter telomere length genetic risk scores and disease incidence in IPF ('narrow'), IPF ('broad'), IPF males & IPF females (using the sex adjusted genetic risk scores) and COPD for genetic instruments derived from 7 and 6 variants. Results were adjusted for registration age, principal ancestral components and assessment centre and also for sex in mixed sex groups.

Figures

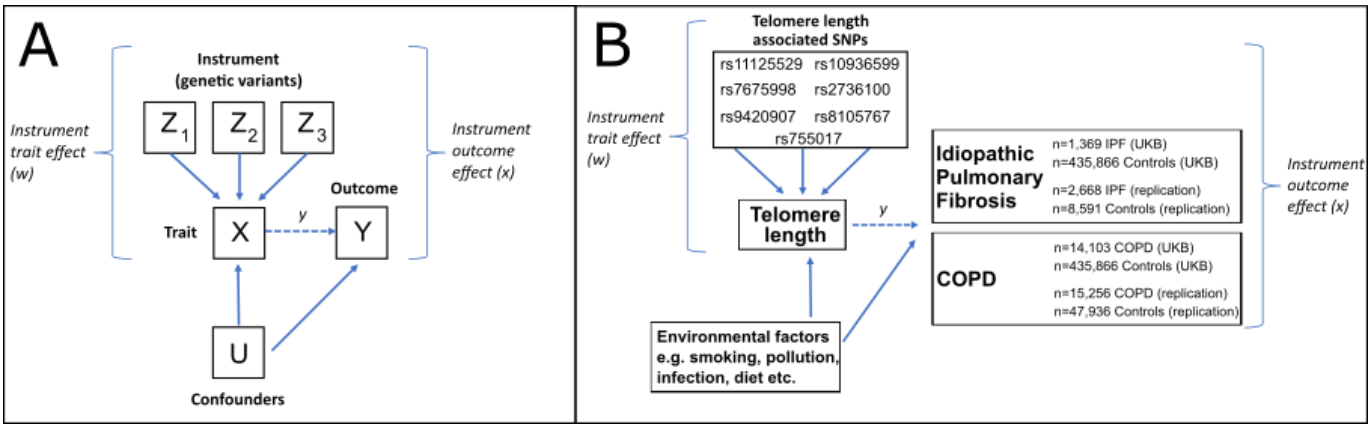

Figure 1. Principle of Mendelian Randomisation; if telomere length plays a causal role in IPF or COPD, genetic variants associated with telomere length will also be associated with the disease. A) If the observed trait X causes the particular outcome Y, the instrument Z (genetic variants associated with the trait) will also be associated with the outcome. **B)** If telomere length causes idiopathic pulmonary fibrosis (IPF) or chronic obstructive pulmonary disease (COPD), the chosen genetic variants associated with telomere length will also be associated with IPF or COPD. As genotype is assigned at conception, it should not be associated with risk factors that might confound the association between telomere length and disease outcome (e.g. smoking or pollution). Weighted estimates of the genetic telomere length association (w) and the genetic disease association (x) are used to infer the causal effect of telomere length on IPF or COPD (y-x/w), which is expected to be free from confounding.

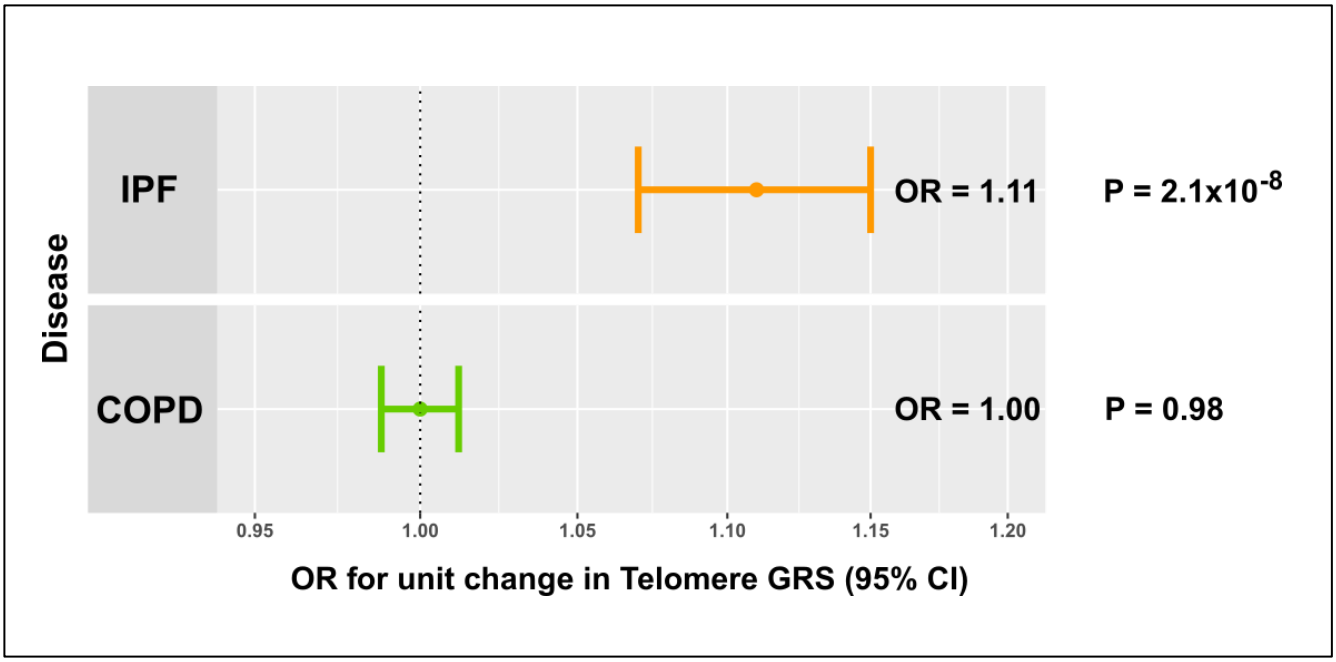

Figure 2. Logistic regression results showing significant disease risk odds ratio for IPF ('narrow') but not for COPD cases for a unit change in telomere length genetic risk score compared with controls in UK Biobank. An odds ratio of 1 shows that disease odds are not influenced by telomere length. An odds ratio significantly different to 1 indicates that odds of disease are influenced by telomere length.

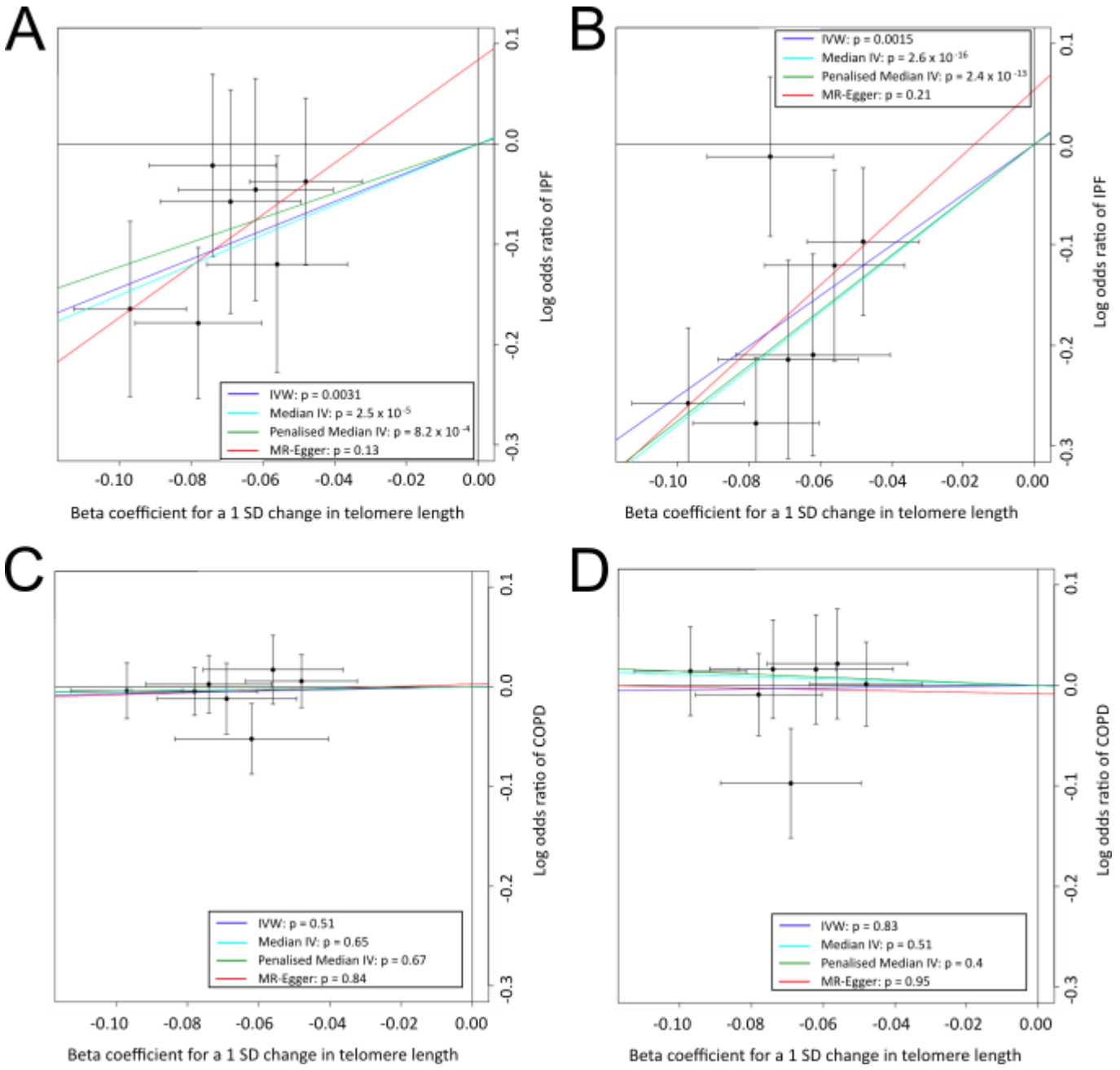

Figure 3. Two-sample MR results for IPF and COPD showing evidence of telomere length causality in IPF but not in COPD. Graphs show the strength of the relationship between disease incidence and telomere length SNPs on the y axis against the telomere length association from previous GWAS for each SNP on the x axis. A non-zero gradient to the lines, with significant p values shown in the top left-hand box for the different MR models used, is evidence of causality of telomere length for disease. Results shown are for all seven telomere variants for (A) IPF ('narrow') in UK Biobank, (B) IPF Replication Cohort, (C) COPD in UK Biobank, (D) COPD Replication Cohort

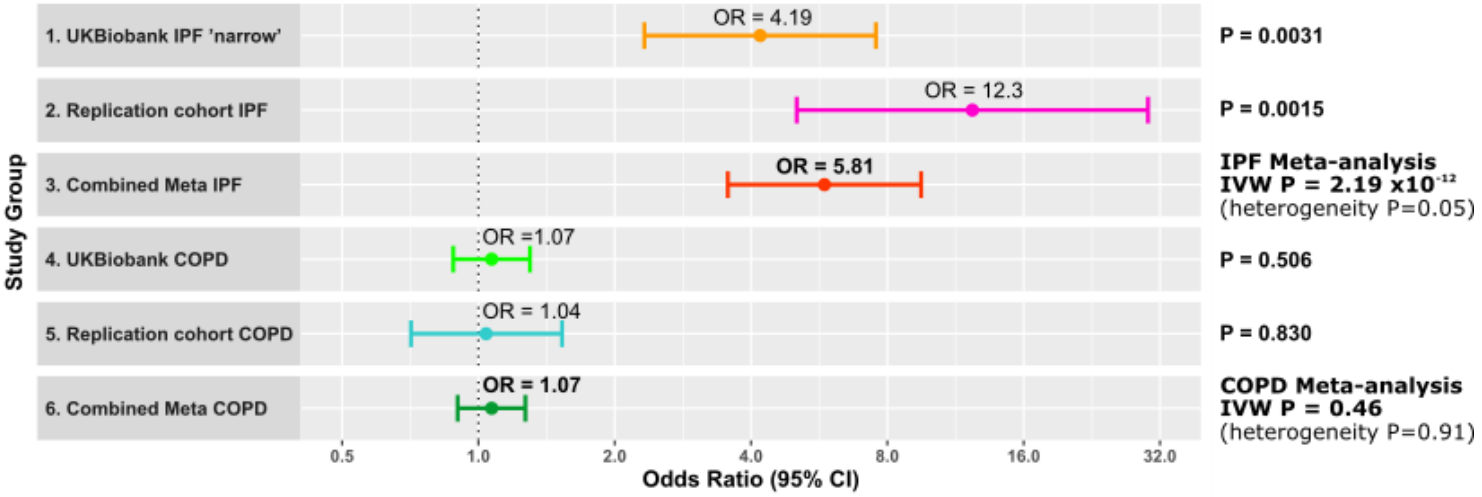

Figure 4. Meta-analysis results for IPF and COPD in UK Biobank and replication cohorts showing significant evidence of telomere length causality in IPF but not COPD across cohorts. Odds ratios and 95% confidence intervals are shown for a genetically-instrumented one standard deviation shorter telomere length in IPF ('narrow') in UKB, IPF Replication Cohort, IPF meta-analysis, COPD in UK Biobank, COPD Replication Cohort and COPD meta-analysis using IVW Method.