Vitamin D status and chronic obstructive pulmonary disease risk: a prospective UK Biobank study

Zheng Zhu,1 Xinglin Wan,2 Jiannan Liu,3 Dandan Zhang,3 Pengfei Luo,1 Wencong Du,1 Lulu Chen,1 Jian Su,1 Dong Hang,2 Jinyi Zhou,2 Xikang Fan1

ABSTRACT

Background Low vitamin D status has been linked to an increased risk for various inflammatory diseases. Conflicting results have been reported regarding chronic obstructive pulmonary disease (COPD). This study aims to investigate the associations of serum 25-hydroxyvitamin D (25(OH)D) concentrations with COPD risk and survival.

Methods We included 403,648 participants with serum 25(OH)D measurements and free of COPD at baseline from UK Biobank. Follow-up was until 30 September 2021. Multivariable-adjusted Cox regression models were applied to estimate HRs and 95% CIs for the associations of season-standardised 25(OH)D concentrations with COPD risk and survival. The restricted cubic splines were used to assess dose–response relationship. Kaplan-Meier estimation was used to create graphs of the survival curves.

Results During a median follow-up of 12.3 (IQR: 11.4–13.2) years, 11,008 cases of COPD were recorded. We observed a non-linear inverse association between 25(OH)D concentrations and COPD risk. Compared with participants in the fourth quintile of 25(OH)D, those in the lowest quintile were associated with a 23% higher risk (HR, 1.23; 95% CI, 1.16 to 1.31). Stronger associations were observed for the risk in men and current smokers (both p for interaction <0.05). In survival analyses, compared with the fourth quintile, cases in the lowest quintile had a 36% higher risk for overall death (HR, 1.38; 95% CI, 1.22 to 1.56).

Conclusion Our findings indicate that serum 25(OH)D concentrations are non-linearly negatively associated with incidence and mortality of COPD, suggesting a potential protective role of vitamin D in the pathogenesis of COPD.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a significant cause of disease burden worldwide.1 The pathological changes observed in COPD cases include chronic inflammation and structural changes due to repeated injury and repair.2 Abundant laboratory research has shown that vitamin D is essential in the acquired and innate immune systems.3 Meanwhile, observational studies have related vitamin D deficiency to a higher risk of infectious, autoimmune and inflammatory diseases.4 5 Although vitamin D status is traditionally known for its effect on skeletal health, existing studies suggest vitamin D status assessed by serum 25-hydroxyvitamin D (25(OH)D) has been related to the regulation of pathogenic mechanisms in COPD.5 6 However, epidemiological evidence regarding the associations of 25(OH)D concentrations with COPD incidence and survival remains inconclusive.

Cross-sectional studies have observed that vitamin D deficiency is highly prevalent among patients with COPD.7 Previous evidence has also associated lower 25(OH)D concentrations with worse lung function and a higher prevalence of COPD.8 10 As patients with COPD mostly exhibit increased skin ageing related to smoking behaviour and reduced sunlight exposure due to less outdoor activity, reductions in 25(OH)D concentrations are considered to be correlated with COPD development.11 Prospective cohort studies associating baseline 25(OH)D with declined lung function or incident COPD have yielded inconsistent results.12–14 Furthermore, results...
from the general population in the Copenhagen City Heart Study indicated an association between lower 25(OH)D concentrations and higher COPD risk,15 but another Danish study using national register data found no such association.16

Although vitamin D deficiency has been associated with higher mortality in the general population,17 18 the effect on patients with COPD remains unclear. Some studies based on patients with COPD found no relationship between 25(OH)D and mortality.19 20 However, post-diagnostic blood vitamin D status in these studies could be primarily modified by COPD induced systemic inflammatory responses.11 Thus, it seems more stable to investigate the association using pre-diagnostic 25(OH)D in patients with COPD.

Therefore, leveraging the serum vitamin D measurements in the UK Biobank, we aimed to assess the associations of 25(OH)D concentrations with COPD incidence in 403 648 participants and survival in 11 008 COPD cases.

METHODS

Patient and public involvement
This study was conducted on UK Biobank database and patients or public were not involved in the development of study design or recruitment of participants.

Study population
UK Biobank is a large prospective cohort study that recruited over 500 000 participants aged 40–69 years in 22 assessment centres across the UK from 2006 to 2010. At the baseline recruitment visit, individuals who signed consent completed a questionnaire assessment and underwent physical measurements. Blood samples were collected from all participants at recruitment and from a subset of participants with a repeat assessment. All participants were registered with the UK National Health Service (NHS). The present study was conducted under the UK Biobank application number 84525.

We excluded participants with airway obstruction (n=49 733), including prevalent COPD at recruitment (n=3516), self-reported emphysema/chronic bronchitis (n=8349), forced expiratory volume at 1 s/forced vital capacity (FEV1/FVC) <lower limit of normal (calculated according to the Global Lung Function Initiative 2012 criteria25) (n=41 779) and participants with unavailable measurements on 25(OH)D (n=49 032). Totally, 403 648 participants were included in the study (online supplemental figure S1).

Assessment of serum 25(OH)D
As part of the UKB Biomarker Project, serum 25(OH)D concentrations were measured by a direct competitive chemiluminescent immunoassay (DiaSorin Liaison XL), with a detection range of 10–375 nmol/L. The mean coefficients of variation were 5.04%, 5.39% and 6.14% for low, medium and high internal quality control levels of 25(OH)D. To further verify the accuracy, the assay of serum 25(OH)D was registered with an external quality assurance (EQA) scheme (RIQAS Immunoassay Speciality 1). According to EQA results, 100% of participated distributions (n=108) were considered good or acceptable.

Ascertainment of COPD
Incident COPD data were available through linkage to the NHS register until 30 September 2021 for England, 31 July 2021 for Scotland and 28 February 2018 for Wales. Cause and date of death after COPD diagnosis were obtained from death certificates until 31 October 2021 for Scotland and 30 September 2021 for England and Wales. The registers are coded according to the 10th revision of the International Classification of Diseases (ICD-10). We used ICD-10 codes to define COPD as J40–44.

Ascertainment of covariates
Covariates were collected from baseline survey and selected according to possible effects on serum 25(OH)D and COPD, including sociodemographic characteristics (age, sex, ethnicity, education and Townsend Deprivation Index), lifestyles (smoking status, alcohol consumption, physical activity and dietary intake), environmental pollution (passive smoking and fine particulate matter, PM2.5), body mass index (BMI) (defined as the weight (kg) divided by height (m) squared), medical history (prevalent asthma and family history of respiratory system diseases) and female-specific factors (menopausal status and hormone replacement therapy). Physical activity was measured as metabolic equivalent task hours per week for moderate activity. PM2.5 estimated for the year 2010 was modelled for each address using a land use regression model. Asthma at baseline was defined using self-report and hospital admission data (ICD-10 codes J45–46). Further details on covariate measurements were recorded in the UK Biobank online protocol.22

Statistical analysis
We calculated season-standardised 25(OH)D concentrations by fitting a cosinor model.25 Baseline characteristics were described among participants grouped by quintiles of 25(OH)D concentrations. The intraclass correlation coefficient (ICC) in two measurements of 25(OH)D between the baseline and the repeated assessments was computed.

In the analysis of COPD risk, participants contributed person-time from the date of recruitment until the date of the initial diagnosis of COPD, lost to follow-up, death or the end of the study, whichever occurred first. Cox proportional hazard models were used to estimate HRs and 95% CIs for the associations of 25(OH)D with COPD risk among total participants, never-smokers and smokers. The proportional hazards assumption was tested using Schoenfeld residuals and no evidence
of non-proportionality was detected. We used multivariate cubic regression splines with four knots to visually explore non-linear associations between serum 25(OH)D concentrations and the risk of COPD.

We classified serum 25(OH)D into quintiles and estimated the HRs using the fourth quintile reflected the typical vitamin D level in UK population as the reference in the primary analysis. We considered three analytic models. Model 1 was adjusted for age, sex, ethnicity and assessment centre. Model 2 was additionally adjusted for education degree, Townsend deprivation index, BMI, smoking status, alcohol drinking, physical activity, oily fish intake, regular use of vitamin D supplements, prevalent asthma, family history of respiratory diseases and female-specific factors. Model 3 was further adjusted for passive smoking, PM2.5 and occupations at risk of COPD. For covariates with missing data, we used the missing indicator method for categorical variables and used the median imputation by gender for continuous variables.

In addition, stratified analyses were conducted by age (≤60, >60 years), sex (male, female), BMI category (<25, 25–30, ≥30 kg/m²), smoking status (never, former and current smokers), passive smoking (never, ever), physical activity (≤median, >median) and use of vitamin D supplements (No, Yes) in the model 3. A likelihood ratio test was performed to examine possible effect modification between 25(OH)D concentrations and the stratification variables, comparing the models with and without interaction terms. Sensitivity analyses were conducted by excluding COPD cases diagnosed within 2 years after baseline assessment, excluding the participants with overall poor self-rated health in baseline questionnaires, and excluding the participants with use of vitamin D supplements.

In survival analyses, we calculated overall and COPD-specific survival time from the date of first diagnosis of COPD to the date of death or end of follow-up. Kaplan-Meier estimation was used to create the graphs of survival curves and the log-rank test was used to compare curves from different quintiles of 25(OH)D. We also use multivariate cox regression to estimate HRs and 95% CIs for overall and COPD-specific survival. The survival analyses were adjusted for the same covariates as the previous analyses.

All analyses were performed using SAS V.9.4 (SAS Institute). Statistical tests were all two-sided and p<0.05 was considered statistically significant.

RESULTS
Among 403 648 participants (185 840 men, 217 808 women) followed for a median of 12.3 years (IQR, 11.4–13.2 years), we documented 11 008 COPD cases (5956 men and 5052 women). Of these COPD cases, 2773 deaths (225 from COPD) occurred within a median survival time of 3.8 years (IQR, 1.6–6.6 years). The ICC in serum 25(OH)D between the initial and repeated assessment was 0.59 (95% CI, 0.58 to 0.60).

The baseline characteristics of participants by quintiles of serum 25(OH)D concentrations are shown in table 1. Participants with higher 25(OH)D had higher economic standards, proportion of oily fish intake, and tended to use vitamin D supplements. The proportion of current smokers and prevalent asthma cases was lower among these participants.

Figure 1 shows a non-linear inverse relationship between 25(OH)D concentrations and COPD incidence among total participants, never-smokers and smokers (all p values for non-linearity <0.0001). We visually observed that the risk of COPD decreased with increasing 25(OH)D concentrations and was lowest at 55 nmol/L in total participants. Table 2 shows the associations between 25(OH)D and COPD risk among total participants, never-smokers and smokers. In the fully adjusted model, compared with the fourth quintile, the lowest quintile groups increase a 23% risk (HR, 1.23; 95% CI, 1.16 to 1.31) in total participants, a 25% risk (HR, 1.25; 95% CI, 1.10 to 1.42) in never-smokers and a 25% risk (HR, 1.23; 95% CI, 1.15 to 1.32) in smokers.

The stratified analysis (online supplemental figure S2) showed that the above associations for COPD generally remained consistent across subgroups. Effect modification by sex and smoking status was observed for COPD risk (both p values for interaction <0.05), and the HRs were stronger in men and current smokers.

Sensitivity analyses showed the results were unchanged after excluding 822 COPD cases diagnosed within 2 years prior to follow-up (online supplemental table S1), excluding 15 695 participants who self-rated overall poor health at baseline assessments (online supplemental table S2) or excluding 98 637 participants with use of vitamin D supplements (online supplemental table S3).

In survival analyses, we observed an association of higher pre-diagnostic 25(OH)D concentrations with a significant increase in overall and COPD-specific survival as analysed by a log-rank test (both p values <0.05) (figure 2). In the fully adjusted models, compared with the fourth quintile of 25(OH)D concentrations, cases in the lowest quintile had a 38% higher risk for overall death (HR, 1.38; 95% CI, 1.22 to 1.56) and 57% higher risk for COPD-specific death (HR, 1.57; 95% CI, 1.03 to 2.40) (table 3).

DISCUSSION
In this large prospective study based on European population, we observed non-linear inverse associations between serum 25(OH)D concentrations and the risk of COPD. The decreasing risk of COPD appeared to be lowest at 55 nmol/L of 25(OH)D. The robustness of the associations was supported by comprehensive stratified and sensitivity analyses. In addition, we also found that lower pre-diagnostic 25(OH)D concentrations were associated with a significant decrease in overall and COPD-specific survival. Our findings imply that vitamin D might play a role in progression of COPD.
Although a growing body of evidence reported that vitamin D deficiency is associated with multiple diseases, the association between vitamin D status and COPD risk remained inconsistent.\(^4\) Cross-sectional and case-control studies indicated that vitamin D deficiency was highly prevalent in patients with COPD.\(^25\)\(^26\) Of note, the prior studies indicated that vitamin D deficiency was highly prevalent in patients with COPD.\(^25\)\(^26\) Of note, the prior studies indicated that vitamin D deficiency was highly prevalent in patients with COPD.\(^25\)\(^26\) However, some other studies found no associations.\(^14\)\(^28\) Such inconsistency may be attributed to smaller sample size, short follow-up period and residual confounding caused by the incomprehensive adjustment. In addition, multiple randomised controlled trials (RCTs) had demonstrated that vitamin D supplementation reduced the rate of worsening moderate or severe COPD exacerbations in patients with low baseline 25(OH)D concentrations.\(^29\)\(^30\)

### Table 1 Baseline characteristics of participants according to quintiles of serum 25(OH)D concentrations

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Serum 25(OH)D concentrations (nmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quintile 1</td>
</tr>
<tr>
<td></td>
<td>&lt;31.7</td>
</tr>
<tr>
<td>n</td>
<td>80704</td>
</tr>
<tr>
<td>Age, mean (SD), years</td>
<td>54.89 (8.13)</td>
</tr>
<tr>
<td>Female, No. (%)</td>
<td>45238 (56.05)</td>
</tr>
<tr>
<td>White race, No. (%)</td>
<td>68748 (85.19)</td>
</tr>
<tr>
<td>College or university degree, No. (%)</td>
<td>29603 (36.68)</td>
</tr>
<tr>
<td>Townsend Index, mean (SD)</td>
<td>-0.49 (3.39)</td>
</tr>
<tr>
<td>Smoking status, No. (%)</td>
<td>Never: 44545 (55.20)</td>
</tr>
<tr>
<td></td>
<td>Previous: 24441 (30.28)</td>
</tr>
<tr>
<td></td>
<td>Current, pack-years &lt;10: 1427 (1.77)</td>
</tr>
<tr>
<td></td>
<td>Current, pack-years ≥10 and &lt;20: 2015 (2.50)</td>
</tr>
<tr>
<td></td>
<td>Current, pack-years ≥20 and &lt;30: 2031 (2.52)</td>
</tr>
<tr>
<td></td>
<td>Current, pack-years ≥30: 3681 (4.56)</td>
</tr>
<tr>
<td>Passive smoking, No. (%)</td>
<td>Never: 48198 (59.72)</td>
</tr>
<tr>
<td></td>
<td>&lt;20 hours a week: 15294 (18.95)</td>
</tr>
<tr>
<td></td>
<td>≥20 hours a week: 1147 (1.42)</td>
</tr>
<tr>
<td>PM(_{2.5}), mean (SD), μg/m(^3)</td>
<td>10.18 (1.11)</td>
</tr>
<tr>
<td>Alcohol intake frequency, No. (%)</td>
<td>Never or special occasions only: 23230 (28.78)</td>
</tr>
<tr>
<td></td>
<td>Once a month to two times a week: 28263 (35.02)</td>
</tr>
<tr>
<td></td>
<td>Three times a week to daily: 28839 (35.73)</td>
</tr>
<tr>
<td></td>
<td>Oily fish intake, No. (%): 67218 (83.30)</td>
</tr>
<tr>
<td></td>
<td>Regular vitamin D supplements, No. (%)</td>
</tr>
<tr>
<td></td>
<td>BMI, mean (SD), kg/m(^2): 28.71 (5.68)</td>
</tr>
<tr>
<td></td>
<td>Physical activity, mean (SD), MET-hour/week: 12.35 (18.46)</td>
</tr>
<tr>
<td></td>
<td>Family history of respiratory disease, No. (%)</td>
</tr>
<tr>
<td></td>
<td>Prevalent asthma, No. (%)</td>
</tr>
<tr>
<td></td>
<td>Occupations at risk of COPD, No. (%)</td>
</tr>
<tr>
<td></td>
<td>Had menopause, No. (%)</td>
</tr>
<tr>
<td></td>
<td>Ever used HRT, No. (%)</td>
</tr>
</tbody>
</table>

*The percentage of the categorical variables does not sum to 100% because some participants chose to ‘not answer’.

BMI, body mass index; COPD, chronic obstructive pulmonary disease; HRT, hormone-replacement therapy; MET, metabolic equivalent; 25(OH)D, 25-hydroxyvitamin D; PM\(_{2.5}\), fine particulate matter.
Furthermore, a meta-analysis based on vitamin D binding protein gene polymorphisms studies also suggested that vitamin D status was associated with COPD risk. In the present study, we found that lower 25(OH)D concentrations could be related to higher COPD risk, providing further evidence for the association between 25(OH)D and COPD.

It is well-established that smoking is the most commonly encountered risk factor for COPD. Our findings indicated that 25(OH)D concentrations were inversely associated with COPD risk in both smokers and never-smokers. Extensive studies had reported smokers had a greater decline in FEV1, a higher incidence of abnormal lung function and a higher mortality rate from COPD. A nested case–control study found a relationship between vitamin D deficiency and COPD in smokers, but lacked evidence in never-smokers. Moreover, a cross-sectional study observed individuals exposed to...
to secondhand smoking have increased risks of COPD.36 The current study showed that the association remained unchanged in never-smokers after adjustment for passive smoking. Furthermore, we also found that lower 25(OH)D concentrations might have a more significant effect on elevating the risk of COPD in current smokers. A meta-analysis including 24 studies with 11 340 participants suggested that smokers were likely to have lower 25(OH)D concentrations.37 Future clinical or laboratory studies are needed to explore the underlying mechanism.

Moreover, there is still debate on the optimal range of 25(OH)D. US Institute of Medicine considered 25(OH)D concentrations of 50 nmol/L (20 ng/mL) or higher to be adequate for bone and overall health,38 while the Endocrine Society stated that of 75 nmol/L (30 ng/mL) or higher provide increased health benefits.39 Our previous finding indicated that 45–60 nmol/L of 25(OH)D might be a potential interventional target for premature death.18 However, the direct evidence regarding optimal 25(OH)D on respiratory diseases was scarce. A recent Mendelian randomisation study suggested that 50–75 nmol/L of 25(OH)D may contribute to lower mortality of respiratory diseases.40 In our analysis, optimal concentrations of 50–60 nmol/L were observed to be associated with lower COPD risk. For COPD survival, our results are consistent with previous studies indicating a negative association between pre-diagnostic 25(OH)D concentrations and COPD mortality.41 42 However, other longitudinal studies found no relationships between post-diagnostic 25(OH)D and COPD survival.12 19 These studies were restricted by the

Figure 2  Survival curves for overall survival (A) and COPD-specific survival (B) according to quintiles of serum 25(OH)D concentrations. 25(OH)D, 25-hydroxyvitamin D; COPD, chronic obstructive pulmonary disease.

Table 3  Associations of serum 25(OH)D concentrations with all-cause or COPD-specific death in COPD cases

<table>
<thead>
<tr>
<th>Models*</th>
<th>Serum 25(OH)D concentrations (nmol/L), HR (95% CI)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>p for non-linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quintile 1</td>
<td>Quintile 2</td>
<td>Quintile 3</td>
<td>Quintile 4</td>
<td>Quintile 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;31.7</td>
<td>31.7 to &lt;41.6</td>
<td>41.6 to &lt;51.8</td>
<td>51.8 to &lt;64.6</td>
<td>≥64.6</td>
<td></td>
</tr>
<tr>
<td>All-cause death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death cases/</td>
<td>884/12 657</td>
<td>536/9874</td>
<td>464/8594</td>
<td>438/8328</td>
<td>451/8551</td>
<td></td>
</tr>
<tr>
<td>person-years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.55 (1.38 to 1.74)</td>
<td>1.12 (0.99 to 1.27)</td>
<td>1.05 (0.92 to 1.20)</td>
<td>Ref</td>
<td>1.00 (0.88 to 1.14)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.38 (1.22 to 1.55)</td>
<td>1.06 (0.93 to 1.20)</td>
<td>1.02 (0.90 to 1.17)</td>
<td>Ref</td>
<td>1.00 (0.88 to 1.14)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.38 (1.22 to 1.56)</td>
<td>1.06 (0.93 to 1.20)</td>
<td>1.02 (0.90 to 1.17)</td>
<td>Ref</td>
<td>1.00 (0.87 to 1.14)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>COPD-specific death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death cases/</td>
<td>82/12 657</td>
<td>56/9874</td>
<td>23/8594</td>
<td>34/8328</td>
<td>30/8551</td>
<td></td>
</tr>
<tr>
<td>person-years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>1.95 (1.30 to 2.93)</td>
<td>1.52 (0.99 to 2.33)</td>
<td>0.67 (0.40 to 1.15)</td>
<td>Ref</td>
<td>0.85 (0.52 to 1.39)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 2</td>
<td>1.59 (1.04 to 2.42)</td>
<td>1.39 (0.90 to 2.14)</td>
<td>0.65 (0.38 to 1.12)</td>
<td>Ref</td>
<td>0.83 (0.51 to 1.36)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.57 (1.03 to 2.40)</td>
<td>1.38 (0.89 to 2.13)</td>
<td>0.64 (0.38 to 1.10)</td>
<td>Ref</td>
<td>0.83 (0.51 to 1.36)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*Model 1: adjusted for age, sex, ethnicity, assessment centre. Model 2: adjusted for model 1 plus college or university degree, Townsend deprivation index, body mass index, smoking status, smoking pack-years, alcohol drinking, summed metabolic equivalent of task-minutes per week for moderate activity, oily fish intake, regular use of vitamin D or multivitamin supplements, prevalent asthma, family history of respiratory diseases and for women, menopause status and hormone replacement therapy. Model 3: adjusted for model 2 plus passive smoking, PM2.5 and occupations at risk of COPD. COPD, chronic obstructive pulmonary disease; 25(OH)D, 25-hydroxyvitamin D; Ref, reference.
small sample size (n=500), and post-diagnostic 25(OH)D concentrations might lead to reverse causation.

Our analysis has several strengths. First, the large sample size and the morbidity and mortality data based on the NHS records provided sufficient power to detect associations. Second, serum 25(OH)D was assessed in a central laboratory using standardised, validated blood biochemistry methods with strict quality control, thereby minimising any measurement errors. Third, we comprehensively adjusted for broad demographics, lifestyle, and medical history. Several limitations also need to be acknowledged. First, reverse causality could not be excluded. However, all participants with baseline COPD were excluded from the analysis, and sensitivity analyses support the robustness of our findings. Second, a single measurement of 25(OH)D at recruitment might not represent long-term exposure concentrations. However, existing studies suggested a single measurement could adequately reflect vitamin D status over an extended period, and the moderate ICC revealed that time-dependent variation in vitamin D was unlikely to produce a substantial influence on our findings. Third, blood samples were collected from all participants at recruitment to minimise selection bias. However, 20% of participants were excluded, with 10% excluded due to missing available measurement for 25(OH)D and the remaining 10% excluded due to airway obstruction at baseline. Finally, since the majority of participants in this study were from the UK, the findings from this study might not apply to other populations.

In conclusion, we observed serum 25(OH)D concentrations were non-linearly associated with COPD risk in both smokers and never-smokers. Our study also indicated that higher concentrations of 25(OH)D were associated with improved survival of COPD. Whether lower concentrations of 25(OH)D are causal or contributory to COPD risk may spur future long-duration and large-scale RCTs.

Acknowledgements We are grateful to UK Biobank participants. This research has been conducted using the UK Biobank resource under application number 84525.

Contributors XF as the guarantor accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish. JZ and XF contributed equally to this paper and are joint corresponding authors. ZZ and PL are joint first authors. XF as the guarantor accepts full responsibility for the work and/or the conduct of the study. ZZ and PL gave full access to all the data in the study and design of the study. LC, PL, WD, DZ and JL critically revised the manuscript for important intellectual content. All authors reviewed and approved the final manuscript.

Funding This work was supported by grants from National Natural Science Foundation of China (No. 81470273) and grants from Clinical Medicine Science and Technology Special Project of Jiangsu Province (No. BL2014083). Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by The UK Biobank study was approved by the North West Multicentre Research Ethics Committee (06/MRE08/65). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available from the authors upon reasonable request and with the permission of the UK Biobank. Requests to access the datasets should be directed to https://www.ukbiobank.ac.uk/, via access@ukbiobank.ac.uk.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES


of cohort studies from Europe and the United States BMJ. BMJ 2014;348:g1766.


UK Biobank, UK Biobank: protocol for a large-scale prospective Epidemiological resource. Available: https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf [Accessed Nov 2022].


Table S1 Sensitivity analyses excluding incident COPD cases within the first two years of follow-up (n = 822) for the associations between serum 25(OH)D and risk of COPD

<table>
<thead>
<tr>
<th>Model</th>
<th>Serum 25(OH)D concentrations (nmol/L), HR (95% CI)</th>
<th>P for non-linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quintile 1 &lt;31.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 2 31.7 to &lt;41.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 3 41.6 to &lt;51.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 4 51.8 to &lt;64.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 5 ≥64.6</td>
<td></td>
</tr>
<tr>
<td>Total participants</td>
<td>No. of COPD/person-years</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Model³</td>
<td>2,654/948,735</td>
<td>1.21 (1.14-1.29)</td>
</tr>
<tr>
<td>Model³</td>
<td>2,021/952,001</td>
<td>1.05 (0.99-1.12)</td>
</tr>
<tr>
<td>Model³</td>
<td>1,829/948,479</td>
<td>1.00 (0.94-1.07)</td>
</tr>
<tr>
<td>Model³</td>
<td>1,809/948,840</td>
<td>ref</td>
</tr>
<tr>
<td>Model³</td>
<td>1,873/946,848</td>
<td>1.03 (0.97-1.10)</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>No. of COPD/person-years</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Model³</td>
<td>565/533,374</td>
<td>1.25 (1.09-1.42)</td>
</tr>
<tr>
<td>Model³</td>
<td>443/544,611</td>
<td>0.99 (0.87-1.13)</td>
</tr>
<tr>
<td>Model³</td>
<td>437/543,143</td>
<td>0.99 (0.87-1.13)</td>
</tr>
<tr>
<td>Model³</td>
<td>446/539,749</td>
<td>ref</td>
</tr>
<tr>
<td>Model³</td>
<td>436/526,886</td>
<td>1.01 (0.89-1.15)</td>
</tr>
<tr>
<td>Smokers</td>
<td>No. of COPD/person-years</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Model³</td>
<td>2,089/415,360</td>
<td>1.20 (1.12-1.29)</td>
</tr>
<tr>
<td>Model³</td>
<td>1,578/407,389</td>
<td>1.07 (0.99-1.15)</td>
</tr>
<tr>
<td>Model³</td>
<td>1,392/405,336</td>
<td>1.01 (0.93-1.08)</td>
</tr>
<tr>
<td>Model³</td>
<td>1,363/409,091</td>
<td>ref</td>
</tr>
<tr>
<td>Model³</td>
<td>1,437/419,962</td>
<td>1.04 (0.96-1.12)</td>
</tr>
</tbody>
</table>

Abbreviations: 25(OH)D = 25-hydroxyvitamin D; COPD = chronic obstructive pulmonary disease; HR = hazard ratio; CI = confidence interval; ref = reference.

³: Model3: adjusted for age, sex, ethnicity, assessment centre, college or university degree, Townsend deprivation index, body mass index, smoking status, smoking pack-years, alcohol drinking, summed metabolic equivalent of task-minutes per week for moderate activity, oily fish intake, regular use of vitamin D or multivitamin supplements, family history of respiratory diseases and for women, menopause status and hormone replacement therapy, prevalent asthma, passive smoking, PM2.5, occupations at risk of COPD.
Table S2 Sensitivity analyses excluding participants with poor self-rated health (n = 15,695) for the associations between serum 25(OH)D and risk of COPD

<table>
<thead>
<tr>
<th>Model</th>
<th>Serum 25(OH)D concentrations (nmol/L), HR (95% CI)</th>
<th>P for non-linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quintile 1 (&lt;31.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 2 (31.7 to &lt;41.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 3 (41.6 to &lt;51.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 4 (51.8 to &lt;64.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quintile 5 (≥64.6)</td>
<td></td>
</tr>
<tr>
<td>Total participants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of COPD/person-years</td>
<td>2,246/88,672</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,859/921,336</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,726/921,336</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,720/926,670</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,792/925,465</td>
<td></td>
</tr>
<tr>
<td>Model3a</td>
<td>1.20 (1.12-1.28)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>No. of COPD/person-years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>472/504,554</td>
<td></td>
</tr>
<tr>
<td></td>
<td>409/528,162</td>
<td></td>
</tr>
<tr>
<td></td>
<td>412/530,369</td>
<td></td>
</tr>
<tr>
<td></td>
<td>431/529,533</td>
<td></td>
</tr>
<tr>
<td></td>
<td>425/516,970</td>
<td></td>
</tr>
<tr>
<td>Model3a</td>
<td>1.19 (1.04-1.37)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Smokers</td>
<td>No. of COPD/person-years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,774/382,171</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,450/388,710</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,314/390,966</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,289/397,137</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,367/408,485</td>
<td></td>
</tr>
<tr>
<td>Model3a</td>
<td>1.20 (1.11-1.29)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Abbreviations: 25(OH)D = 25-hydroxyvitamin D; COPD = chronic obstructive pulmonary disease; HR = hazard ratio; CI = confidence interval; ref = reference.

Model3a: adjusted for age, sex, ethnicity, assessment centre, college or university degree, Townsend deprivation index, body mass index, smoking status, smoking pack-years, alcohol drinking, summed metabolic equivalent of task-minutes per week for moderate activity, oily fish intake, regular use of vitamin D or multivitamin supplements, family history of respiratory diseases and for women, menopause status and hormone replacement therapy, prevalent asthma, passive smoking, PM2.5, occupations at risk of COPD.
Table S3 Sensitivity analyses excluding participants with regular use of vitamin D supplements (n = 98,637) for the associations between serum 25(OH)D and risk of COPD

<table>
<thead>
<tr>
<th>Model</th>
<th>Serum 25(OH)D concentrations (nmol/L), HR (95% CI)</th>
<th>P for non-linearity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quintile 1</td>
<td>Quintile 2</td>
</tr>
<tr>
<td></td>
<td>&lt;31.7</td>
<td>31.7 to &lt;41.6</td>
</tr>
<tr>
<td>Total participants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of COPD/person-years</td>
<td>2,562/818,287</td>
<td>1,819/776,544</td>
</tr>
<tr>
<td>Model3a</td>
<td>1.22 (1.14-1.30)</td>
<td>1.04 (0.97-1.12)</td>
</tr>
<tr>
<td>Non-smokers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of COPD/person-years</td>
<td>516/458,584</td>
<td>387/443,306</td>
</tr>
<tr>
<td>Model3a</td>
<td>1.26 (1.09-1.46)</td>
<td>1.02 (0.88-1.19)</td>
</tr>
<tr>
<td>Smokers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of COPD/person-years</td>
<td>2,046/359,704</td>
<td>1,432/333,239</td>
</tr>
<tr>
<td>Model3a</td>
<td>1.20 (1.11-1.30)</td>
<td>1.05 (0.97-1.14)</td>
</tr>
</tbody>
</table>

Abbreviations: 25(OH)D = 25-hydroxyvitamin D; COPD = chronic obstructive pulmonary disease; HR = hazard ratio; CI = confidence interval; ref = reference.

Model3: adjusted for age, sex, ethnicity, assessment centre, college or university degree, Townsend deprivation index, body mass index, smoking status, smoking pack-years, alcohol drinking, summed metabolic equivalent of task-minutes per week for moderate activity, oily fish intake, family history of respiratory diseases and for women, menopause status and hormone replacement therapy, prevalent asthma, passive smoking, PM2.5, occupations at risk of COPD.
UK Biobank baseline participants [n = 502 413]

Participants who were free of COPD at baseline [n = 452 680]

Participants included in the analysis [n = 403 648]

Incident COPD cases during follow-up [n = 11 008]

Sensitivity analysis

All-cause deaths [n = 2 773]
COPD deaths [n = 225]

Participants with airway obstruction (n = 49 733):
- Prevalent COPD (n = 3 516)
- Self-reported emphysema/chronic bronchitis (n = 8 349)
- FEV₁/FVC < LLN (n = 41 779)

Participants who had missing serum 25(OH)D measurements (n = 49 032)

Excluding cases within two years prior to follow-up (n = 822)

Excluding participants who self-rated poor health (n = 15 695)

Excluding participants used vitamin D supplements (n = 98 637)

Figure S1 Flow chart of study participants
Abbreviations: COPD = chronic obstructive pulmonary disease; FEV₁ = forced expiratory volume in the first second; FVC = forced vital capacity; LLN = lower limit of normal.
<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Quintile</th>
<th>HR (95% CI)</th>
<th>P for interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60</td>
<td>Q1</td>
<td>1.28 (1.14-1.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.04 (0.83-1.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.98 (0.77-1.21)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>0.87 (0.65-1.16)</td>
<td></td>
</tr>
<tr>
<td>≥80</td>
<td>Q1</td>
<td>1.24 (1.15-1.34)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.09 (0.92-1.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.03 (0.87-1.21)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.05 (0.88-1.25)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Q1</td>
<td>1.35 (1.24-1.48)</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.10 (1.01-1.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.04 (0.95-1.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>0.96 (0.80-1.15)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Q1</td>
<td>1.20 (1.01-1.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.06 (0.90-1.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.96 (0.80-1.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.01 (0.86-1.19)</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;19</td>
<td>Q1</td>
<td>1.26 (1.11-1.42)</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.17 (1.03-1.33)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.05 (0.91-1.22)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>0.98 (0.84-1.13)</td>
<td></td>
</tr>
<tr>
<td>25-30</td>
<td>Q1</td>
<td>1.24 (1.15-1.34)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.06 (0.90-1.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.98 (0.83-1.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.01 (0.86-1.19)</td>
<td></td>
</tr>
<tr>
<td>≥30</td>
<td>Q1</td>
<td>1.26 (1.10-1.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.01 (0.86-1.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.02 (0.87-1.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.10 (0.96-1.26)</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>Q1</td>
<td>1.25 (1.10-1.42)</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>0.99 (0.87-1.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.99 (0.87-1.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.02 (0.89-1.17)</td>
<td></td>
</tr>
<tr>
<td>Previous</td>
<td>Q1</td>
<td>1.17 (1.07-1.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.08 (0.95-1.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.99 (0.88-1.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.03 (0.90-1.18)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>Q1</td>
<td>1.38 (1.23-1.56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.16 (1.03-1.31)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.07 (0.94-1.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.03 (0.90-1.18)</td>
<td></td>
</tr>
<tr>
<td>Passive smoking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>Q1</td>
<td>1.17 (1.07-1.28)</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.02 (0.93-1.13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.96 (0.88-1.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.03 (0.94-1.13)</td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>Q1</td>
<td>1.28 (1.14-1.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.11 (0.98-1.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.05 (0.94-1.17)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.02 (0.91-1.14)</td>
<td></td>
</tr>
<tr>
<td>Physical activity (MET-hour/week)</td>
<td>Q1</td>
<td>1.25 (1.17-1.34)</td>
<td>0.99</td>
</tr>
<tr>
<td>Below median</td>
<td>Q2</td>
<td>1.12 (1.01-1.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.02 (0.91-1.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.10 (0.89-1.35)</td>
<td></td>
</tr>
<tr>
<td>Above median</td>
<td>Q1</td>
<td>1.27 (1.14-1.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.05 (0.94-1.16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.02 (0.89-1.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.10 (0.89-1.34)</td>
<td></td>
</tr>
<tr>
<td>Vitamin D supplement</td>
<td>No</td>
<td>1.24 (1.16-1.33)</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.05 (0.96-1.17)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>0.97 (0.88-1.08)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>0.96 (0.87-1.06)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Q1</td>
<td>1.30 (1.13-1.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>1.14 (0.98-1.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>1.13 (0.99-1.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>1.13 (0.99-1.28)</td>
<td></td>
</tr>
</tbody>
</table>

Figure S2 Forest plots of stratified analysis of the association between serum 25(OH)D concentrations and COPD risk
Abbreviations: COPD = chronic obstructive pulmonary disease; BMI = body mass index; MET = metabolic equivalent; HR = hazard ratio; CI = confidence interval; ref = reference.