BMJ Open Respiratory Research

Higher serum vitamin D levels are associated with decreased odds of obstructive lung disease in the general population: an NHANES analysis (2007–2008 to 2009–2010)

Mohamed Ismail Seedahmed ,[1] Aaron D Baugh,[1] Jordan A Kempker[2]

To cite: Seedahmed MI, Baugh AD, Kempker JA. Higher serum vitamin D levels are associated with decreased odds of obstructive lung disease in the general population: an NHANES analysis (2007–2008 to 2009–2010). *BMJ Open Resp Res* 2020;**7**:e000798. doi:10.1136/bmjresp-2020-000798

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2020-000798).

Received 12 October 2020
Revised 13 December 2020
Accepted 14 December 2020

[1]Pulmonary, University of California San Francisco, San Francisco, California, USA
[2]Pulmonar, Emory University, Atlanta, Georgia, USA

Correspondence to
Dr Mohamed Ismail Seedahmed; Mohamed.Seedahmed@ucsf.edu

ABSTRACT

Background Obstructive lung disease is a significant cause of morbidity and healthcare burden within the USA. A growing body of evidence has suggested that vitamin D levels can influence the course or incidence of obstructive lung disease. However, there is an insufficient previous investigation of this association.

Study design and methods We used the National Health and Nutrition Examination Survey (NHANES) cycles 2007–2008 and 2009–2010 spirometry results of individuals aged 40 years and older to assess the association between serum 25-hydroxyvitamin D levels and obstructive lung disease, as defined by the American Thoracic Society using the lower limit of normal. We used stage multivariate survey-logistic regression.

Results The final model included age, gender, body mass index, pack-years smoking history, season, income-to-poverty ratio and race/ethnicity. In the primary analysis using vitamin D as a continuous variable, there was no association between vitamin D levels and obstructive lung disease. We noted a trend between 'other Hispanic' self-identified race and serum vitamin D levels wherein higher levels were associated with higher odds of obstructive lung disease in this ethnicity, but not among other racial or ethnic groups (OR (95% CI)=1.40 (0.98 to 1.99), p=0.06). In a secondary analysis, when vitamin D was measured as a categorical variable, there was a significant association between the highest levels of serum vitamin D levels and lesser odds of obstructive lung disease (OR (95% CI)=0.77 [0.61 to 0.98], p=0.04).

Conclusions Higher serum vitamin D levels among adults are associated with decreased odds of obstructive lung disease in the general population. Results among non-Mexican Hispanic participants highlight the need for further research in minority populations. More work is needed to address the course and incidence of lung disease in the USA.

Key messages

- In the general population, is there an independent association between vitamin D and obstructive lung disease after controlling for relevant covariates?
- Higher serum vitamin D levels are independently associated with decreased odds of obstructive lung disease in the general US population.
- This paper adds nuance to the broad understanding of vitamin D's role in lung pathophysiology.

INTRODUCTION

Obstructive airway disease is a significant public health concern in the USA. Chronic Obstructive Lung Disease (COPD) was estimated to cost 16.4 million lost workdays in 2010 and has greater odds of prompting an end to employment than diabetes or heart disease.[1 2] In the next 20 years, asthma is projected to cost over US$900 billion to the US economy.[3] Together with COPD, it is the fourth leading cause of death in the USA.[4] Given the enormity of these challenges, there is an urgent need to identify interventions that can reduce obstructive airway disease incidence and burden.

While vitamin D has been classically described regarding healthy bone metabolism, there has been increased attention towards its extra-skeletal physiological actions in recent years. Low levels of 25-hydroxyvitamin D (25(OH)D) have been linked to clinical conditions such as rickets, hypertension, ischaemic heart disease, diabetes mellitus type 1, some cancers, osteoporosis and infections.[5] Its role in lung development and pathophysiology has also received increasing attention. Gestational deficiency is negatively associated with later pulmonary function and confers increased odds of asthma.[6 7] In adults, vitamin D deficiency has been linked to an increased likelihood of respiratory diseases.[8] Besides, many of the demographic groups with a higher incidence of vitamin D deficiency[9] also report more significant morbidity from obstructive airway diseases.[10] Data from the

National Health and Nutrition Examination Survey (NHANES) 2005–2006 revealed that 41.6% of adult participants ≥20 years old had vitamin D deficiency, with higher prevalence among non-Hispanic blacks (82.1%), those with no college education and those with body mass index (BMI) more than 30.[9] Also, data from NHANES 2001–2004 demonstrated that older age, female sex, winter season and smoking are associated with vitamin D deficiency.[11]

However, much of the adult literature has been reported in diseased populations. In these populations, associations can be highly confounded. vitamin D levels are influenced by sun exposure and, therefore, maybe a proxy for disability.[12] Further exploring the relationship between vitamin D and pulmonary health in adults would benefit from extensive, robust studies of the general population to understand this association and its mediators. From there, the goal of this study was to use a large, representative sample of US adults to examine the relationship between vitamin D status and obstructive lung disease patterns among US adults aged 40 years and older.

METHODS

Data source and study design

NHANES is a major programme of the National Center for Health Statistics (NCHS) which is part of the Centers for Disease Control and Prevention. It is an ongoing national survey designed to assess the health of the general US population. Data are collected annually on a 2-year cycle, using a multistage, probability-sampling design to generate population-level estimates. The national survey employs a design variable in order to approximate the civilian, non-institutionalised US population. NHANES 2007–2010 included spirometry examinations conducted according to the technical recommendations of the American Thoracic Society (ATS) for procedures and equipment.[13] We included all participants from the 2007–2008 and 2009–2010 survey cycles who were at least 40 years of age, completed spirometry, and had measured serum 25(OH)D concentrations (figure 1). Per their study protocols, NHANES excluded participants with BMI >40, a history of tuberculosis; supplemental oxygen usage; a history of haemoptysis or retinal detachment on attempted spirometry; or any recent stroke, active cardiovascular disease, retinal, thoracic or abdominal surgery.

Public-use data from NHANES were obtained from files available on the NCHS website.[14] The data were sorted, merged and concatenated using the unique sequence number given to each NHANES participants, in addition to a specific identifier number that code for the 2-year cycle, 2007–2008 and 2009–2010.[15]

Patient involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Exposure variable

The NHANES 2007–2010 used ultra-high-performance liquid chromatography-tandem mass spectrometry for all vitamin D measurements.[13] For the first analysis, we created a new continuous vitamin D variable, divided by 25, for which a change in 1 unit will equal a change in 25 nmol/L of vitamin D. As a secondary analysis, we considered vitamin D as a categorical variable and tried to account for the ongoing uncertainties around this question. We found the evidence marshalled by the Institute of Medicine around 30 nmol/L as a cut-off for deficiency in the general population convincing, and adopted it as one threshold.[16] However, we also recognised that this threshold was adopted with regards to bone health, pulmonary outcomes, having never received systematic consideration. Diverging from the Institute of Medicine, the Endocrine Society's guidelines focused on 'high-risk' individuals and recommended a higher cut-off of 72.5 nmol/L.[17] Several obstructive lung diseases were listed within this category, and without weighing in on their appropriateness with regards to musculoskeletal disease, we thought this a reasonable approach to capture the otherwise unknown pulmonary risk. Similarly, we also observed that the Chair of the relevant committee at the Institute of Medicine had opined that 75 nmol/L represented a likely ceiling for beneficial effects in the literature.[16] Thus, our synthesis of available evidence led us to adopt 75 nmol/L as a second cut-off to create a three-level categorical variable.

Outcome variable

The primary outcome variable in our analysis was the presence or absence of obstructive lung disease. We employed the ATS definition of obstructive disease as having a ratio of forced expiratory volume in 1 s to forced vital capacity (FEV_1/FVC), which is less than the fifth percentile lower limit of normal (LLN) observed in the healthy, non-smoking population.[18 19] Per contemporaneous ATS guidelines, the spirometric measurements were recoded to calculate the LLN using reference equations developed in 1999 from participants in NHANES III.[18 20]

Additional covariates

Smoking is recognised as the most crucial risk factor for the development of obstructive lung disease.[21] Low BMI is associated with increased COPD risk and mortality.[22] Both BMI and pack-years smoking history were measured as continuous variables. Because seasonality can affect performance on spirometry,[23] we included a 6-month interval binary variable to control for the time of a participant's testing. Age, gender, race and income have important associations with vitamin D levels in the USA and were all measured as covariates.[24] Race/ethnicity was reported categorically according to the US Census classifications. Income was measured using the income-to-poverty ratio(IPR) calculated as a ratio of the reported

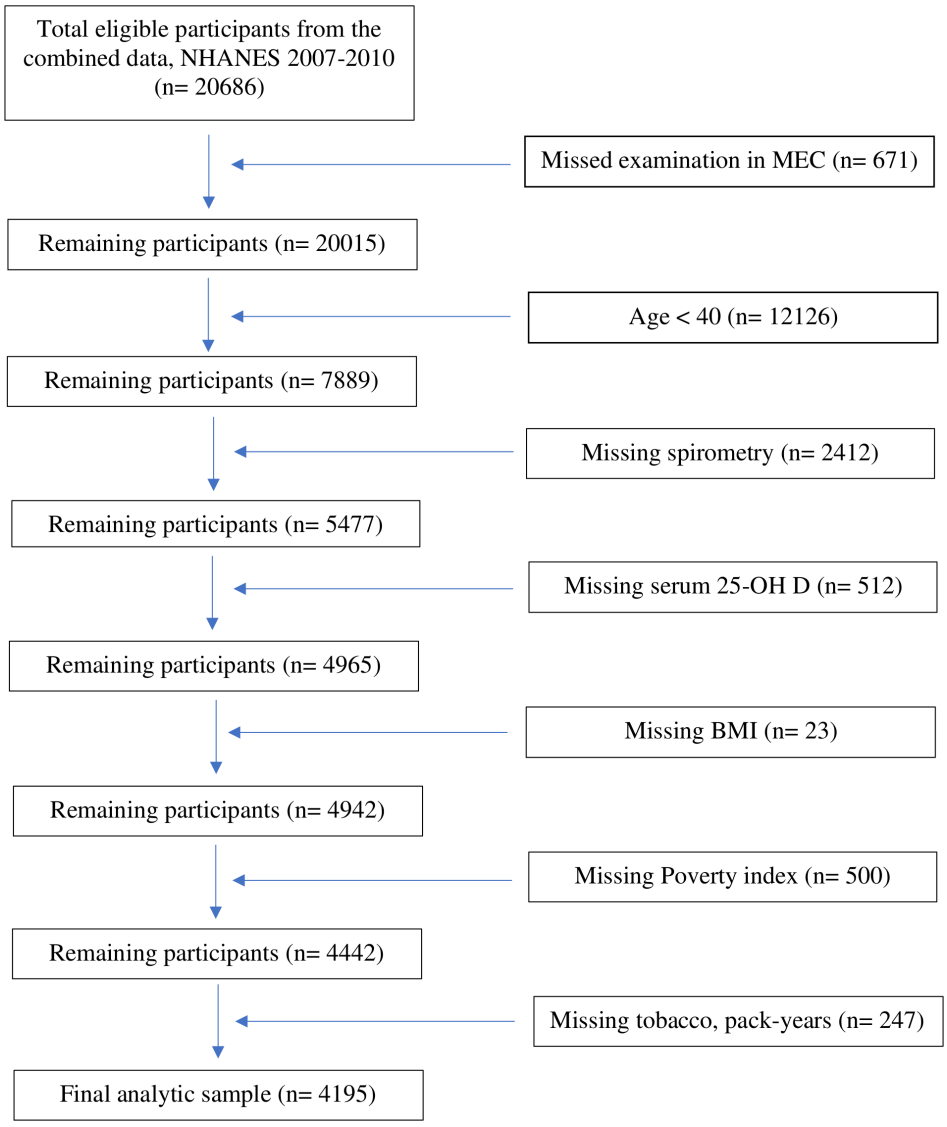

Figure 1 Strengthening the Reporting of Observational Studies in Epidemiology flow chart-sample selection criteria for the association between serum 25-hydroxyvitamin D (25(OH)D) concentration and baseline forced expiratory volume in 1 s/forced vital capacity. BMI, body mass index; NHANES, National Survey of the National Center for Health Statistics; MEC, mobile examination centres; 25(OH)D, total vitamin D.

household income to the national poverty threshold of the given year reported by the US Census Bureau.[25]

Analyses

All statistical analyses were performed with SAS software (V.9.4). Descriptive statistics were computed. For all analyses, we used the NHANES 2007–2010 sample weights previously calculated for the combined two survey cycles considered.[26] The PROC SURVEY procedure was employed to account for the intricate sampling design of NHANES. We followed the Taylor Series linearisation method and made an assumption of non-random missingness for variance estimation, per NHANES analytic guidelines.[27] We also used domain analysis to refine our variance estimate.

In our primary analysis, we developed an a priori multivariate survey logistic regression model of the association between obstructive lung disease, defined as FEV_1/FVC<LLN, and serum 25(OH)D concentration as a continuous variable. Variables were preselected based on identifying major determinants of obstructive lung disease in the previous medical literature. Following a prespecified multistep approach, multivariate models were adjusted for age, gender, race/ethnicity, smoking in pack-years, BMI, the season of examination and income-to-poverty ratio (IPR). The modelling building strategy started with assessing for collinearity.[28] We assessed for interaction terms between serum 25(OH)D and other covariates following this step. Next, we conducted a confounding assessment. The prevalence OR of the a priori model was compared with all confounders' possible subsets using the 10% rule (online supplemental file). For all model estimates, we defined significance as non-overlapping 95% CIs and p≤0.05. While our analysis

did not identify any statistical evidence of confounding in this dataset. The weight of evidence in previous literature and the structural limitations in data collection led us to retrain the a priori model as our final, which better captured the real-world determinants of lung function. Moreover, we performed a subgroup analysis for the association between serum 25(OH)D as continuous and obstructive lung disease stratified by race/ethnicity. Finally, to clarify policy implication and real-world interpretation, we conducted a secondary analysis of vitamin D as a categorical variable. A planned tertiary analysis by the Global Initiative for Chronic Obstructive Lung Disease stages was not completed due to the small number of participants with stage 3 and 4 disease.

To help understand the differences between each race/ethnicity subgroups among subjects with obstructive lung disease defined by LLN, we computed the mean of post-bronchodilator FEV_1 (% predicted) and serum 25(OH)D for each race/ethnicity category. Additionally, we used the ellipse statement to graphically plot the predicted ellipses for each race/ethnicity subgroups.

RESULTS

Of the 20 686 participants from NHANES 2007–2010, 20 015 had conducted the survey and examinations in the mobile examination centres. Four thousand one hundred and ninety-five met the inclusion criteria for our study (figure 1).

Study participants' characteristics as related to the baseline FEV_1/FVC

Fourteen and half of a per cent (n=777) of the study population met the diagnostic criteria for obstructive lung disease with a baseline FEV_1/FVC<LLN (table 1). They were more likely non-Hispanic Whites (61% vs 47%, p≤0.0001), greater than 10 pack-year total lifetime smokers (55% vs 26%, p≤0.0001), to have a higher mean of serum 25(OH)D concentration (71±1.14, p≤0.0001), and to have a higher age mean (56.1 vs 54.7, p≤0.0001). Compared with participants with FEV_1/FVC≥LLN, they were more likely Hispanics (29% vs 16%, p≤0.0001), and likely to be obese with BMI ≥30 (42% vs 27%, p≤0001). There were no significant differences between the two groups, FEV_1/FVC below and above LLN, by gender, PIR or seasonality of examination administration.

Study participants' characteristics as related to the 25(OH)D status

Those with vitamin D deficiency constituted only 8% (n=569) of our sample. In comparison to those with adequate serum vitamin D measurements, these participants had higher odds of being female gender (58% vs 50.2%, p=0.001), non-Hispanic Black race/ethnicity (52% vs 15%, p≤0.0001), obese with BMI ≥30 (50% vs 39%, p≤0.0001), and higher poverty-index-ratio <1 (78% vs 82%, p≤0.0001), greater than 10 pack-year total lifetime smokers (38% vs 31, p=0.02), examined between 1 November and 30 April (62% vs 43%, p≤0001).

Association between obstructive lung disease and serum 25(OH)D

In the final multivariate-adjusted model (a priori model), no statistically significant association was appreciated between obstructive lung disease and serum 25(OH)D as a continuous variable (OR (95% CI)=0.96 (0.86 to 1.07), p=0.46) (table 2). In a secondary analysis of the final model (a priori model) using vitamin D as a categorical variable, there was a significant association between higher vitamin D levels and decreased odds of obstructive lung disease (OR (95% CI)=0.77 (0.61 to 0.98), p=0.04) (table 2).

While interaction testing did not reveal a significant interaction between obstructive lung disease and self-reported race/ethnicity, there was a strong trend towards significance. To further explore this point, we performed a subgroup analysis of the final model stratified by race/ethnicity. Results revealed that, while not associated with obstructive lung disease in non-Hispanic Whites, serum 25(OH)D levels showed a trend towards statistically significant association with obstructive lung disease in other (non-Mexican) Hispanics (OR (95% CI)=1.40 (0.98 to 1.99), p=0.06) (table 2). Among those with obstructive lung disease below the LLN, other (non-Mexican) Hispanics had a lower mean of serum 25(OH)D (65.4 nmol/L, 95% CI 60.1 to 70.8), compared with non-Hispanic Whites (mean=74.7, 95% CI 72.0 to 77.4) (figure 2). The mean of percentage predicted post-bronchodilator FEV_1 was slightly higher in other (non-Mexican) Hispanics (mean=89.8 %, 95% CI 85.5 to 94.4), compared with non-Hispanic Whites (mean=85.7, 95% CI 83.0 to 88.5) (online supplemental file). Additionally, the 95% prediction non-Mexican Hispanic's ellipse is slightly thinner than other race subgroups, indicating that the correlation between baseline FEV_1 (% predicted) and total vitamin D is greater among non-Mexican Hispanics (figure 3).

DISCUSSION

Using the NHANES 2007–2010 data, we explored an association between obstructive lung disease and serum 25(OH)D level in the general population. In our final model, vitamin D status was not independently associated with a diagnosis of obstructive lung disease when measured as a continuous variable, but higher levels were associated with lower odds when measured as a categorical variable. When stratified by race/ethnicity, there was a trend towards a positive association between serum vitamin D levels and odds of obstructive lung disease among non-Mexican Hispanics.

The most similar study to our own, the population-based examination by Ganji *et al*, failed to detect any association between vitamin D levels and obstructive lung disease.[29] In using ATS definitions rather than self-report,

Table 1 Demographics and clinical characteristics of study participants by baseline FEV_1/FVC ratio below and above LLN,* NHANES, 2007–2010

	Baseline FEV_1/FVC<LLN†		Baseline FEV_1/FVC≥LLN†		
	N	%	N	%	
Characteristic	**777**	**14.5**	**4700**	**85.5**	**P value‡**
Age, years					
40–49	195	25	1459	31	<0.0001
50–59	202	26	1270	27	
60–69	210	27	1219	26	
≥70	170	22	752	16	
Mean±SE	56.1±0.41		54.7±0.22		<0.0001
Race					
Mexican American	72	9	852	18	<0.0001
Other Hispanic	55	7	536	11	
Non-Hispanic White	474	61	2230	47	
Non-Hispanic Black	154	20	892	20	
Other race or multiracial	22	3	190	4	
Gender					
Male	445	57.3	2280	48.5	0.21
Female	332	42.7	2420	51.5	
BMI, kg/m^2					
<18.5	26	3	33	1	<0.0001
18.5–25	258	33	946	20	
25–30	283	37	1716	37	
≥30	205	27	1987	42	
Missing	5		18		
Income-to-poverty ratio¶					
Below poverty level (<1.0)	130	18	691	16	0.40
Above poverty level (≥1.0)	594	82	3562	84	
Missing	53		447		
Total 25(OH) vitamin D, nmol/L					
<30	49	7	329	8	0.80
30–74	422	60	2607	61	
≥75	229	33	1329	31	
Mean±SE	71±1.14		69.8±0.90		<0.0001
Missing	77		435		
Tobacco, pack-years					
0	203	28	2530	56	<0.0001
1–10	122	17	787	18	
>10	409	55	1179	26	
Missing	43		204		
The 6-month examination period					
1 November–30 April	307	40	2165	46	0.15
1 May–31 October	470	60	2535	54	
Spirometry measurements (mean±SE)					
Baseline FEV_1/FVC in %	61.8±0.33		77.8±0.17		<0.0001

Continued

Table 1 Continued

Characteristic	Baseline FEV_1/FVC<LLN† N	Baseline FEV_1/FVC<LLN† %	Baseline FEV_1/FVC≥LLN† N	Baseline FEV_1/FVC≥LLN† %	P value‡
	777	14.5	4700	85.5	
Post-bronchodilator FEV_1 (% predicted)	390	50	183	4	<0.0001
Missing	85.5±1.13		91.8±1.43		<0.0001
	387		4517		

*Lower limit of normal.
†Column percentage from the total N. The percentage is calculated from the respective column based on baseline FEV_1/FVC.
‡Rao-Scott χ^2 test.
§Missing values and per cent: represent column percent from the total N based on baseline FEV_1/FVC.
¶Income-to-poverty ratio ($income/$threshold)=represents the ratio of family or unrelated individual income to their appropriate poverty threshold. It was calculated by dividing family income by the poverty guidelines, specific to family size, appropriate year and state.
BMI, body mass index; FEV_1, forced expiratory volume in 1 s/forced vital capacity; FVC, forced vital capacity; LLN, lower limit of normal; NHANES, National Health and Nutrition Examination Survey; 25(OH)D, 25-hydroxyvitamin D.

we addressed a major potential weakness of that study. This is most likely responsible for the divergent results. Additionally, our analysis adjusted for several important covariates unconsidered in this earlier analysis. Several previous studies have suggested an association, as we also found. For instance, a recent metanalysis found that as compared with those without COPD, those with the disease had a lower serum vitamin D level.[30] These results also help to rationalise studies suggesting a role for vitamin D in the pathogenesis of obstructive lung disease.[31 32]

Our results accord well with previous findings in diseased populations documenting negative outcomes in association with vitamin D insufficiency.[29] Impressively, this is true even in spite of the generally poor correlation between spirometric airflow obstruction, symptom burden and exacerbation frequency.[21] Though limited by their small sample size, some randomised controlled trials have reported improvements in exacerbations, 6-min walk distance, or symptoms as measured by the COPD assessment test.[33 34] Longitudinal cohorts of COPD populations similarly demonstrate increased exacerbations, accelerated FEV_1 decline and higher symptom burden in association with lower vitamin D levels.[35] Cumulatively, the weight of evidence supports the notion that low serum vitamin D levels may be associated with worse pulmonary health outcomes.

Table 2 The crude and multivariable-adjusted associations between 25(OH)D and baseline FEV_1/FVC<LLN,* NHANES, 2007–2010

Vitamin D measurements		Crude model† OR	Crude model† 95% CI	Crude model† P value	A priori model‡ OR	A priori model‡ 95% CI	A priori model‡ P value
25(OH)D per 25 nmol/L§		1.06	0.96 to 1.16	0.24	0.96	0.86 to 1.07	0.46
Stratified by race/ethnicity							
25(OH)D per 25 nmol/L§	Mexican American	1.19	0.90 to 1.58	0.21	0.89	0.68 to 1.16	0.36
	Other Hispanic	1.46	1.05 to 2.02	0.03	1.40	0.98 to 1.99	0.06
	Non-Hispanic White (Ref.)	0.98	0.88 to 1.10	0.77	0.94	0.84 to 1.06	0.29
	Non-Hispanic Black	0.88	0.70 to 1.10	0.24	0.80	0.61 to 1.05	0.10
	Other race, including multiracial	1.33	0.91 to 1.96	0.14	1.31	0.71 to 2.45	0.38
Total 25(OH)D, nmol/L¶							
<30 (deficiency)		0.88	0.57 to 1.36	0.55	0.89	0.53 to 1.50	0.65
30–74 (insufficient=Ref.)		–	–	–	–	–	–
≥75 (sufficient)		0.96	0.77 to 1.19	0.69	0.77	0.61 to 0.98	0.04

*Lower limit of normal.
†Crude model included only vitamin D measurement only in the model.
‡Final model (a priori model) included the following covariates: vitamin D, age, gender, race/ethnicity, BMI, smoking, Season and income-to-poverty ratio.
§Total vitamin D as continues variables, total vitamin D=total vitamin D/25, 1 unit change equal to a change by 25 nmol/L of vitamin D.
¶Total vitamin D as a categorical variable.
BMI, body mass index; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; LLN, lower limit of normal; NHANES, National Health and Nutrition Examination Survey; 25(OH)D, 25-hydroxyvitamin D.

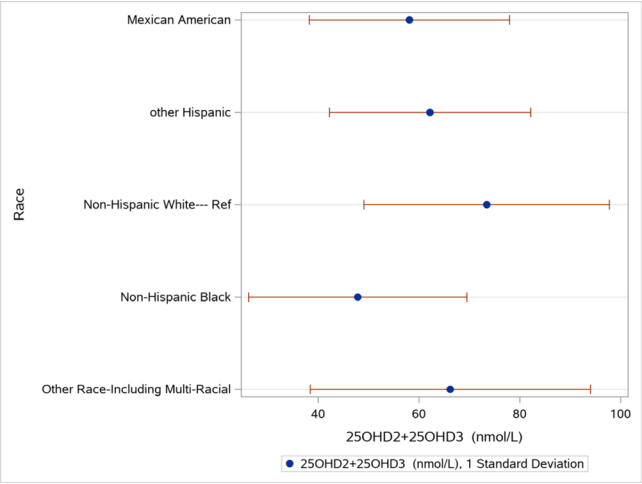

Figure 2 Dot plots of total vitamin D by race/ethnicity among subjects with obstructive lung disease below the lower limit of normal. Each dot represents the mean of serum 25-hydroxyvitamin D (25(OH)D) for each race/ethnicity category, and bands represent the SD. FEV_1, forced vital capacity in 1 s; 25(OH)D2+25(OH)D3, total 25(OH)D.

In our subgroup analysis by self-identified racial and ethnic groups, we observed a trend towards significance in non-Mexican Hispanics. In this group, a possible association between higher levels of vitamin D and obstructive lung disease was noted. Hispanics as a whole are underserved in healthcare and enormously under-represented in clinical trials.[36] Neighbourhood quality or other social and structural determinants of health may be essential in investigating the existence of this association. Hispanic-majority neighbourhoods were burdened with higher environmental toxin exposure.[37 38] Further, there are positive correlations between air pollution exposure and walkability, especially in low-income communities.[39] These effects might plausibly account for higher serum 25(OH)D and higher incidence of obstructive lung disease, respectively. However, in our view, the present dataset was ill-suited to such explorations. Although NHANES data are structured as such, it is far from evident that a Cuban-American whose family was naturalised as refugees some three generations ago is meaningfully similar in circumstances or outlook to an undocumented Honduran migrant. Within this designated subgroup alone, previous literature identified different asthma prevalence,[40] smoking patterns,[41] disease-specific mortality[42] and health insurance coverage.[43] Such heterogeneity precludes meaningful conclusions. These results merit study in an appropriately composed of study population with methodologies well-designed to interrogate this specific question.

Our study had several other important limitations. The cross-sectional design precludes any discussion of causality. While skin complexion affects vitamin D absorption,[44] and African ancestry is associated with lower mean lung function,[45] self-identified race can be a poor proxy for both.[46 47] Comorbid restrictive disease

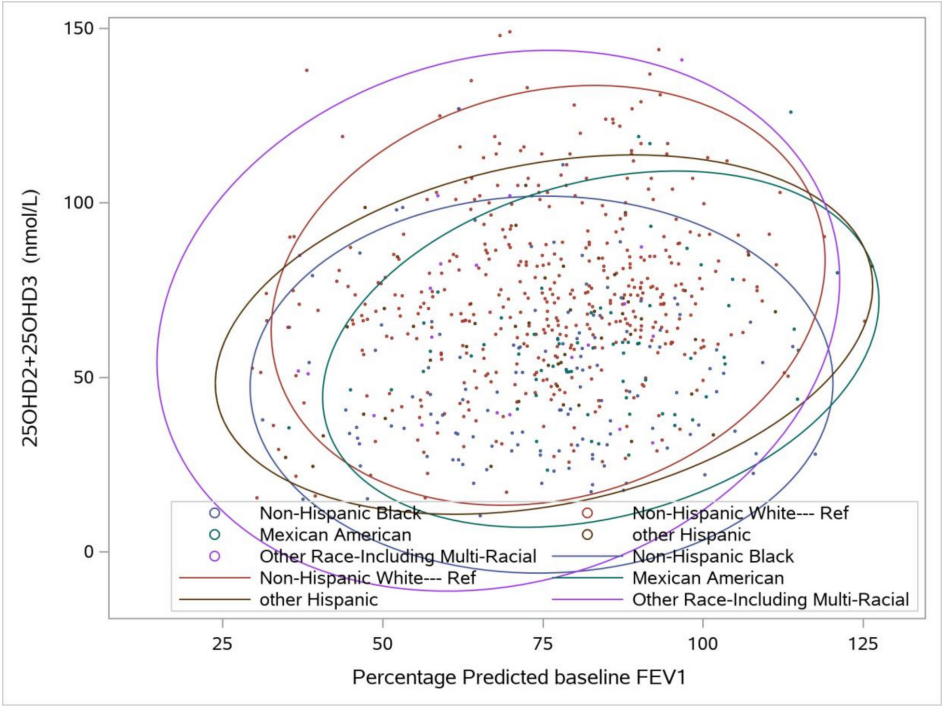

Figure 3 95% prediction ellipses for baseline FEV_1 (% predicted) by each race subgroup among subjects with obstructive lung disease below the LLN. The means of the variables (the centres of the ellipses) are different across the race subgroups. The larger the ellipse, the greater the variance within that race subgroup. FEV_1, forced vital capacity in 1 s; LLN, lower limit of normal; 25(OH)D2+25(OH)D3, total 25-hydroxyvitamin D (25(OH)D).

can obscure underlying obstructive physiology, but the NHANES omitted body plethysmography that would allow for its assessment. Socioeconomic status is a complex, multi-faceted confounder that we measured in only one dimension, only at the individual level and not at all at the neighbourhood level, limiting our appreciation of its effects. Obstructive lung diseases are a broad category with many distinct etiologies. While our data did not allow us to distinguish between them, it is unlikely that all have the same association with serum vitamin D levels, or even that the same pathophysiology is responsible for said associations. Given the known timeline of lung maturation and the critical period for other vitamins, measurements of vitamin D from earlier periods in life would have been more informative than contemporaneous samples.[48 49]

In a population-based study of the USA derived from the NHANES 2007–2008 and 2009–2010 cycles, in the general population, we found an association between higher serum 25(OH)D levels and lesser odds of obstructive lung disease by spirometric criteria. Among non-Mexican Hispanics, there was a trend towards increasing serum 25(OH)D levels associated with increased odds for obstructive lung disease. This finding calls attention to the importance of further research on minority health as the USA grows increasingly diverse. Altering the prevalence of obstructive lung disease in the USA will require a multi-faceted approach, including nutritional, healthcare access and other interventions, which these findings will help to inform.

Correction notice This article has been corrected since it first published. The provenance and peer review statement has been included.

Acknowledgements The authors would like to thank the faculty of the Rollins School of Public Health at Emory University, more specifically, the thesis committee chair, Veronika Fedirko, MPH, PhD, for her tremendous guidance and support with the implementation of the algorithm, study design, data cleaning and analysis. Additionally, the authors would like to thank Dr Mehrdad Arjomandi, Dr Prescott Woodruff and Dr Neeta Thakur for their valuable review of the manuscript.

Contributors Conceived and designed the study research: MIS, JAK. Developed study protocol: MIS, JAK. Worked on the methods: MIS, JAK. Analysed and interpreted data: MIS, ADB, JAK. Prepared the manuscript: MIS, ADB, JAK.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests JAK has previously received grants from the US National Institute of Health and fees from Grifols, both related to work on separate topics of personal interest.

Patient consent for publication Not required.

Ethics approval This study was approved by the NCHS Research Ethics Review Board (ERB).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. The National Health and Nutrition Examination Survey (NHANES) is an ongoing national survey project of the Center for Disease Control and Prevention (CDC) designed to assess the general US population's health. Also, NHANES is a major programme of the National Center for Health Statistics (NCHS), which is part of the CDC and responsible for producing vital and health statistics for the Nation. The survey is unique in that it combines interviews and physical examinations. Data are collected annually on a 2-year cycle, using a multistage, probability-sampling design to generate population-level estimates. In our study, we included all participants from the 2007–2008 and 2009–2010 survey cycles. Information and data are made available, on the NHANES website, to the public and researchers worldwide. NHANES Website https://www.cdc.gov/nchs/nhanes/index.htmNHANES Data Release and Access Policy https://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf NHANES 2007-2008 Data: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2007 NHANES 2009-2010 Data:https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx?BeginYear=2009.

ORCID iD

Mohamed Ismail Seedahmed http://orcid.org/0000-0002-7446-7346

REFERENCES

1 Thornton Snider J, Romley JA, Wong KS, *et al*. The disability burden of COPD. *COPD* 2012;9:513–21.
2 Ford ES, Murphy LB, Khavjou O, *et al*. Total and state-specific medical and absenteeism costs of COPD among adults aged ≥ 18 years in the United States for 2010 and projections through 2020. *Chest* 2015;147:31–45.
3 Yaghoubi M, Adibi A, Safari A, *et al*. The projected economic and health burden of uncontrolled asthma in the United States. *Am J Respir Crit Care Med* 2019;200:1102–12.
4 Kochanek KD, Murphy SL, Xu J. Deaths: final data for 2017. *Natl Vital Stat Rep* 2019;68:1–77.
5 Kempker JA, Magee MJ, Cegielski JP, *et al*. Associations between vitamin D level and hospitalizations with and without an infection in a national cohort of Medicare beneficiaries. *Am J Epidemiol* 2016;183:920–9.
6 Gazibara T, den Dekker HT, de Jongste JC, *et al*. Associations of maternal and fetal 25-hydroxyvitamin D levels with childhood lung function and asthma: the generation R study. *Clin Exp Allergy* 2016;46:337–46.
7 Zosky GR, Berry LJ, Elliot JG, *et al*. Vitamin D deficiency causes deficits in lung function and alters lung structure. *Am J Respir Crit Care Med* 2011;183:1336–43.
8 Hejazi ME, Modarresi-Ghazani F, Entezari-Maleki T. A review of vitamin D effects on common respiratory diseases: asthma, chronic obstructive pulmonary disease, and tuberculosis. *J Res Pharm Pract* 2016;5:7–15.
9 Forrest KYZ, Stuhldreher WL. Prevalence and correlates of vitamin D deficiency in US adults. *Nutr Res* 2011;31:48–54.
10 Han MK, Curran-Everett D, Dransfield MT, *et al*. Racial differences in quality of life in patients with COPD. *Chest* 2011;140:1169–76.
11 Looker AC, Pfeiffer CM, Lacher DA, *et al*. Serum 25-hydroxyvitamin D status of the US population: 1988-1994 compared with 2000-2004. *Am J Clin Nutr* 2008;88:1519–27.
12 Kunisaki KM, Niewoehner DE, Singh RJ, *et al*. Vitamin D status and longitudinal lung function decline in the lung health study. *Eur Respir J* 2011;37:238–43.
13 NHANES questionnaires, datasets, and related documentation. Available: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx [Accessed 16 Sep 2020].
14 CDC - NCHS - National Center for Health Statistics. Available: https://www.cdc.gov/nchs/index.htm [Accessed 12 Oct 2020].
15 NHANES - NCHS Research Ethics Review Board Approval. Available: https://www.cdc.gov/nchs/nhanes/irba98.htm [Accessed 12 Oct 2020].
16 Ross AC, Manson JE, Abrams SA, *et al*. The 2011 report on dietary reference intakes for calcium and vitamin D from the Institute of medicine: what clinicians need to know. *J Clin Endocrinol Metab* 2011;96:53–8.
17 Holick MF, Binkley NC, Bischoff-Ferrari HA, *et al*. Evaluation, treatment, and prevention of vitamin D deficiency: an endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011;96:1911–30.

18 Miller MR, Hankinson J, Brusasco V, *et al*. Standardisation of spirometry. *Eur Respir J* 2005;26:319–38.
19 Culver BH, Graham BL, Coates AL, *et al*. Recommendations for a standardized pulmonary function report. An official American thoracic Society technical statement. *Am J Respir Crit Care Med* 2017;196:1463–72.
20 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159:179–87.
21 Vogelmeier CF, Criner GJ, Martinez FJ, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. gold executive summary. *Am J Respir Crit Care Med* 2017;195:557–82.
22 Park HJ, Cho JH, Kim HJ, *et al*. The effect of low body mass index on the development of chronic obstructive pulmonary disease and mortality. *J Intern Med* 2019;286:573–82.
23 Gratziou C, Rovina N, Lignos M, *et al*. Exhaled nitric oxide in seasonal allergic rhinitis: influence of pollen season and therapy. *Clin Exp Allergy* 2001;31:409–16.
24 Wallace TC, Reider C, Fulgoni VL. Calcium and vitamin D disparities are related to gender, age, race, household income level, and weight classification but not vegetarian status in the United States: analysis of the NHANES 2001-2008 data set. *J Am Coll Nutr* 2013;32:321–30.
25 How the census bureau measures poverty. Available: https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html [Accessed 19 Aug 2020].
26 Afzal S, Lange P, Bojesen SE, *et al*. Plasma 25-hydroxyvitamin D, lung function and risk of chronic obstructive pulmonary disease. *Thorax* 2014;69:24–31.
27 Zipf G, Chiappa M, Porter KS, *et al*. National health and nutrition examination survey: plan and operations, 1999-2010. *Vital Health Stat* 2013;1:1–37.
28 Midi H, Sarkar SK, Rana S. Collinearity diagnostics of binary logistic regression model. *J Interdiscip Math* 2010;13:253–67.
29 Ganji V, Al-Obahi A, Yusuf S, *et al*. Serum vitamin D is associated with improved lung function markers but not with prevalence of asthma, emphysema, and chronic bronchitis. *Sci Rep* 2020;10:11542.
30 Zhu M, Wang T, Wang C, *et al*. The association between vitamin D and COPD risk, severity, and exacerbation: an updated systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis* 2016;11:2597–607.
31 Wolsk HM, Chawes BL, Litonjua AA, *et al*. Prenatal vitamin D supplementation reduces risk of asthma/recurrent wheeze in early childhood: a combined analysis of two randomized controlled trials. *PLoS One* 2017;12:e0186657.
32 Janssens W, Bouillon R, Claes B, *et al*. Vitamin D deficiency is highly prevalent in COPD and correlates with variants in the vitamin D-binding gene. *Thorax* 2010;65:215–20.
33 Rafiq R, Prins HJ, Boersma WG, *et al*. Effects of daily vitamin D supplementation on respiratory muscle strength and physical performance in vitamin D-deficient COPD patients: a pilot trial. *Int J Chron Obstruct Pulmon Dis* 2017;12:2583–92.
34 Alavi Foumani A, Mehrdad M, Jafarinezhad A, *et al*. Impact of vitamin D on spirometry findings and quality of life in patients with chronic obstructive pulmonary disease: a randomized, double-blinded, placebo-controlled clinical trial. *Int J Chron Obstruct Pulmon Dis* 2019;14:1495–501.
35 Burkes RM, Ceppe AS, Doerschuk CM, *et al*. Associations Among 25-Hydroxyvitamin D Levels, Lung Function, and Exacerbation Outcomes in COPD: An Analysis of the SPIROMICS Cohort. *Chest* 2020;157:856–65.
36 Rangel ML, Heredia NI, Reininger B, *et al*. Educating Hispanics about clinical trials and biobanking. *J Cancer Educ* 2019;34:1112–9.
37 Cushing L, Faust J, August LM, *et al*. Racial/Ethnic disparities in cumulative environmental health impacts in California: evidence from a statewide environmental justice screening tool (calenviroscreen 1.1). *Am J Public Health* 2015;105:2341–8.
38 Adkins A, Barillas-Longoria G, Martínez DN, *et al*. Differences in social and physical dimensions of perceived walkability in Mexican American and non-Hispanic white walking environments in Tucson, Arizona. *J Transp Health* 2019;14:100585.
39 James P, Hart JE, Laden F. Neighborhood walkability and particulate air pollution in a nationwide cohort of women. *Environ Res* 2015;142:703–11.
40 Rose D, Mannino DM, Leaderer BP. Asthma prevalence among US adults, 1998-2000: role of Puerto Rican ethnicity and behavioral and geographic factors. *Am J Public Health* 2006;96:880–8.
41 Wainwright K, Perrotte JK, Bibriescas N, *et al*. Smoking expectancies and health perceptions: an analysis of Hispanic subgroups. *Addict Behav* 2019;98:106008.
42 Frankenfield DL, Krishnan SM, Ashby VB, *et al*. Differences in mortality among Mexican-American, Puerto Rican, and Cuban-American dialysis patients in the United States. *Am J Kidney Dis* 2009;53:647–57.
43 Hispanic/Latino - The Office of Minority Health. Available: https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=3&lvlid=64 [Accessed 5 Dec 2020].
44 Clemens TL, Adams JS, Henderson SL, *et al*. Increased skin pigment reduces the capacity of skin to synthesise vitamin D3. *Lancet* 1982;1:74–6.
45 Kumar R, Seibold MA, Aldrich MC, *et al*. Genetic ancestry in lung-function predictions. *N Engl J Med* 2010;363:321–30.
46 Yaeger R, Avila-Bront A, Abdul K, *et al*. Comparing genetic ancestry and self-described race in African Americans born in the United States and in Africa. *Cancer Epidemiol Biomarkers Prev* 2008;17:1329–38.
47 Wilson JF, Weale ME, Smith AC, *et al*. Population genetic structure of variable drug response. *Nat Genet* 2001;29:265–9.
48 Knudson RJ, Lebowitz MD, Holberg CJ, *et al*. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983;127:725–34.
49 Zhang H, Garber SJ, Cui Z, *et al*. The angiogenic factor midkine is regulated by dexamethasone and retinoic acid during alveolarization and in alveolar epithelial cells. *Respir Res* 2009;10:77.