Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

Genetic variation underlying common hereditary hyperbilirubinaemia (Gilbert's syndrome) and respiratory health in the 1946 British birth cohort

Laura J. Horsfall[1,3,*], Rebecca Hardy[2], Andrew Wong[2], Diana Kuh[2], Dallas M. Swallow[3]

[1]*Research Department of Primary Care and Population Health, Institute of Epidemiology & Health, University College London, UK;* [2]*MRC University Unit for Lifelong Health and Ageing at University College London, UK;* [3]*Research Department of Genetic, Environment and Evolution, University College London, UK*

Background & Aims: Bilirubin has potent antioxidant properties *in vitro* and raised serum levels have been associated with lower rates of respiratory disease. The enzyme uridine diphosphate glucuronosyltransferase polypeptide 1A1 (UGT1A1) is solely responsible for clearing bilirubin from the blood and homozygosity for seven thymine-adenine (TA) repeats in the TATA box regulatory element of the *UGT1A1* gene underlies a mild hereditary unconjugated hyperbilirubinaemia (Gilbert's syndrome). Our aim was to investigate whether this genetic variation is associated with differences in respiratory health.

Methods: The relationship between the promoter genotype underlying Gilbert's syndrome (*UGT1A1* rs8175347 [TA]7/7) and respiratory outcomes assessed at ages 43, 53, and 60–64 were examined in 2190 members of the 1946 British birth cohort.

Results: The (TA)7/7 genotype, present in 9% of the cohort, was associated with higher forced expiratory volume (FEV1) and forced vital capacity (FVC). The relationship was strongest for heavy smokers (⩾20 cigarettes per day) at age 53 with mean FEV1 409 ml higher (191 to 627; p <0.001) and mean FVC 530 ml higher (95% CI 262–798; p <0.001) for *UGT1A1* (TA)7/7 Gilbert's syndrome participants than for all others, indicating a protection from the pulmonary consequences of heavy smoking. The odds of respiratory disease (chronic obstructive pulmonary disease, self-reported asthma, or prescription of respiratory drugs) were half in those with Gilbert's syndrome genotype (odds ratio 0.49 [95% CI 0.39–0.74]; p <0.001) compared to those without this genotype.

Conclusions: Genetically raised unconjugated serum bilirubin is associated with higher adult respiratory function and protection from respiratory disease.

Keywords: Bilirubin; Uridine diphosphate glucuronosyltransferase; Chronic obstructive pulmonary disease; Respiratory function; Asthma; Polymorphism; Cohort study.

Received 28 November 2013; received in revised form 17 July 2014; accepted 21 July 2014; available online 31 July 2014

* Corresponding author. Address: Research Department of Primary Care and Population Health, University College London Medical School, Royal Free Campus, London NW3 2PF, UK. Tel.: +44 (0) 207 794 0500; fax: +44 (0) 207 794 1224.

E-mail address: laura.horsfall@ucl.ac.uk (L.J. Horsfall).

Abbreviations: UGT1A1, uridine diphosphate glucuronosyltransferase polypeptide 1A1; TA, thymine-adenine; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; ROS, reactive oxygen species; MRC, Medical Research Council; NSHD, National Survey of Health and Development; COPD, chronic obstructive pulmonary disease.

Introduction

Respiratory systems are exposed to reactive oxygen species (ROS) from sources such as cigarette smoke, air pollution and indoor cooking/heating fires [1]. Interindividual differences in the blood levels of ROS scavengers may partially explain differences in the susceptibility to smoking-related respiratory disease [2].

Bilirubin is the yellow-coloured end product of the haem catabolism pathway in mammals. The enzyme haem oxygenase generates an insoluble form of bilirubin from haem that is transported to the liver and is converted to a water-soluble (conjugated) form by the enzyme uridine diphosphate-glucuronosyltransferase (UGT1A1). Potent cytoprotective properties including antioxidant have been reported for both the unconjugated and conjugated forms of bilirubin [3,4]. Various lines of experimental evidence mainly involving animal models suggest a role for serum bilirubin in protecting respiratory tissues against environmental stressors [5]. For example, bilirubin infusions can protect against bleomycin-induced pulmonary fibrosis in rats [6]. Cohort studies have reported lower rates of respiratory diseases and increased respiratory function in people with comparatively higher serum bilirubin levels after accounting for important confounders including smoking status [7–9]. However reverse causation where the disease process alters bilirubin levels or residual confounding by unmeasured confounders could potentially explain these relationships.

Genome wide association studies have repeatedly identified *UGT1A1* as the major genetic locus underlying bilirubin levels in European, East Asian and African populations [10–12]. The *UGT1A1* rs8175347 thymine-adenine (TA) polymorphism of the TATA box regulatory motif is well characterized in European populations since homozygosity for a low expresser variant ([TA]7 or [TA]8 repeats *vs.* [TA]5 or [TA]6) underlies a common form of

familial unconjugated hyperbilirubinaemia (Gilbert's syndrome; OMIM 143500) [13]. Finding an association between this common promoter variation, causing unconjugated hyperbilirubinaemia, and improved respiratory health could support a protective role for bilirubin previously suggested by animal experiments and observational studies in humans [14]. Unlike disease association studies using serum bilirubin levels, the potential for association being due to reverse causation is low because the *UGT1A1* genotype is determined at birth and unaffected by the disease process. The random allocation of parental alleles at conception should in principle balance observed and unobserved confounders across genotype groups reducing the potential for residual confounding. This type of genetic association study is sometimes termed "Mendelian randomization" although causal inference using this design relies on some important assumptions [14].

In the present study, we used data from the Medical Research Council (MRC) National Survey of Health and Development (NSHD), commonly known as the 1946 British birth cohort study, to investigate the relationship between a genetic variation in *UGT1A1*, underlying Gilbert's syndrome, and adult respiratory health as well as interactions with age and smoking status.

Patients and methods

Study population

The NSHD is the longest continuously running birth cohort in the world with details on sampling strategy, data collections, attrition and representativeness reported elsewhere [15–17]. In brief, the NSHD comprises a socially stratified sample of 5362 births out of a total of 16,695 registered in England, Wales and Scotland in one week in the year 1946 [16,18]. The cohort includes all singleton births to wives of non-manual and agricultural workers and 1 in 4 singleton births to wives of manual workers. The cohort participants have been followed since birth with the most recent contact phase between the ages of 60 and 64 [15]. At age 53, 3035 of the original cohort were invited to provide a DNA sample. Contact was not attempted for individuals who had already died (n = 476), were living abroad (n = 583), had previously refused to take part (n = 648), or were untraced since the last contact at 43 years (n = 266). The sociodemographic features of the participants at ages 43 and 53 were broadly representative of the British post-war generation compared with the 1991 England Census data for the national population but with some over-representation of the widowed, or married women and men with lower educational attainment [17]. The sociodemographic features of participants at ages 60–64 appear to be similar to the 2001 England Census reference population, although there were lower levels of home ownership and limiting illness [15]. The cohort participants were selected prior to major immigration flows and so white British ethnicity is over-represented relative to the UK adult population. Additional details on the cohort, including numbers lost to follow-up at various contact phases with reasons, are included in the Supplementary Patients and methods section. DNA was obtained and successfully extracted from blood and/or buccal samples for 2939 out of the 3035 participants at age 53 [18] The study protocol was approved by the Medical Research and Ethics Committee (approval reference MREC 98/2/121 and MREC 07/H1008/168).

Variant selection and genotyping

The *UGT1A1* rs8175347 (TA)5–8 repeat variant of the TATA box motif and the rs4124874 c.-3279T>G SNP of the phenobarbital enhancer module (PBREM) were selected for testing because these two loci have good evidence of regulatory function and explain around 40–45% of the estimated variation in bilirubin levels in healthy Europeans [19–23]. Although these two loci are in a fairly strong linkage disequilibrium ($r^2 = 0.67$) independent effects of the PBREM SNP over the (TA)n repeat on serum bilirubin levels have been reported [19]. The *UGT1A1* (TA)n repeat was assayed using buccal samples (n = 2939) by a previously reported size-separation technique using high-percentage polyacrylamide gels [24]. The KASPar system by LGC Genomics, UK (http://www.lgcgenomics.com), was used to type the PBREM SNP using blood samples (n = 2718). Call rates were >95% for each variant. The reliability of the genotyping was checked by re-typing the (TA)n repeat and the PBREM SNP in a random subset of duplicate samples. Concordance was over 95% and there was also no evidence of deviation from the Hardy–Weinberg equilibrium ($p > 0.05$). All genotyping was done blind without information on phenotypes.

Outcome definition

Forced expiratory volume in litres in one second (FEV1) and forced vital capacity (FVC) in litres were measured during home visits at age 43 in 1989, at age 53 in 1999 and in clinical research facilities or at home visits between ages 60 and 64 at the most recent follow-up during 2006–2011 using a Micro Medical turbine electronic spirometer (Cardinal Health UK 232 Ltd, Basingstoke, UK). Three measures were recorded at age 43 and two at ages 53 and 60–64. Only participants where at least two respiratory measures at each time point were taken and deemed satisfactory by the nurse were included. The maximum readings at each of these time points were used in the analysis. Of the 2939 participants with DNA, acceptable respiratory function measures were available at age 43 for 2151, at age 53 for 2190 and at age 60–64 for 1531.

The presence of respiratory disease was defined as either having chronic obstructive pulmonary disease (COPD), asthma or a respiratory drug prescription at any of the follow-up time points. Moderate to severe COPD was spirometrically defined as a predicted FEV1 below 80% and a FEV1/FVC ratio below 0.7 (i.e. Global initiative for chronic obstructive lung disease stage 2 or higher). Predicted FEV1 was estimated using published algorithms derived from the British population [25]. Presence of an asthma diagnosis was based on the participant reporting recurring asthma at age 43 or clinically diagnosed asthma at age 53. Reports of asthma were not collected at ages 60–64. Respiratory drug use was defined as a self-reported prescription of bronchodilators, corticosteroids, cromoglycate, oxygen or mucolytics at ages 43, 53 or 60–64 years.

Statistical analyses

Mixed linear regression models, using a random subject effect to account for multiple measures in the same participants, were used to examine the relationship between genotypes and mean respiratory function. The models included a random intercept and random slope for age, specifying an unstructured covariance matrix. We categorized the common genotypes as high expresser group 0 (homozygous TA6), intermediate expresser group 1 (heterozygotes) and low expresser group 2 (homozygous TA7) [22,23]. The rare (TA)5 allele was treated as (TA)6, and (TA)8 as (TA)7 in the genotype categories [22,23]. Based on existing literature, the effect of this locus on bilirubin levels does not conform to the additive genetic model with levels for group 1 only slightly higher than group 0 but approximately double for group 2 *vs.* either group 0 or 1. Thus, we fitted the locus as a three level categorical variable and compared this with a simpler (recessive) model with group 2 *vs.* groups 0/1. The *UGT1A1* PBREM SNP (rs4124874) was categorized as high expressers (homozygous T), intermediate expressers (heterozygotes) and low expressers (homozygous G), based on existing literature and the SNP was included as a three level variable to test for any independent effect on respiratory function over the (TA)n locus. Regression models initially included important predictors of respiratory function, which were age (continuous), sex, smoking status, height (continuous), and region of birth. Further details on how these variables were measured are included in the Supplementary Patients and methods. Although there is no prior evidence to suspect these variables would differ by genotype they are important predictors of the respiratory outcome and as such could improve the precision of estimation for other covariates in the regression models [26]. Previous studies have reported that the importance of genetic influences on respiratory function may become weaker with age [27] and stronger in the presence of ROS exposure [28–30] and thus, in a pre-specified analysis we tested for interactions between *UGT1A1* variation, age and smoking status. If any significant interactions were identified, stratified analyses were performed. Statistical significance of categorical variables in regression models including genotypes and interactions with genotypes were tested using the Wald test. The PBREM SNP was added to the models already containing (TA)n genotypes and the Wald test was used to test whether this locus exhibited an independent effect over (TA)n.

A discrete-time proportional odds survival model was fitted to estimate the risk of respiratory disease across genotypes. The three time periods were included as dummy variables in the model (birth to age 43, age 43 to 53, and age 53 to age 60–64) and a random subject effect was included to account for multiple observation periods in the same participants.

Mendelian randomization is a form of instrumental variable analysis for causal inference using observational data [14,31]. The phenotype of interest (i.e.

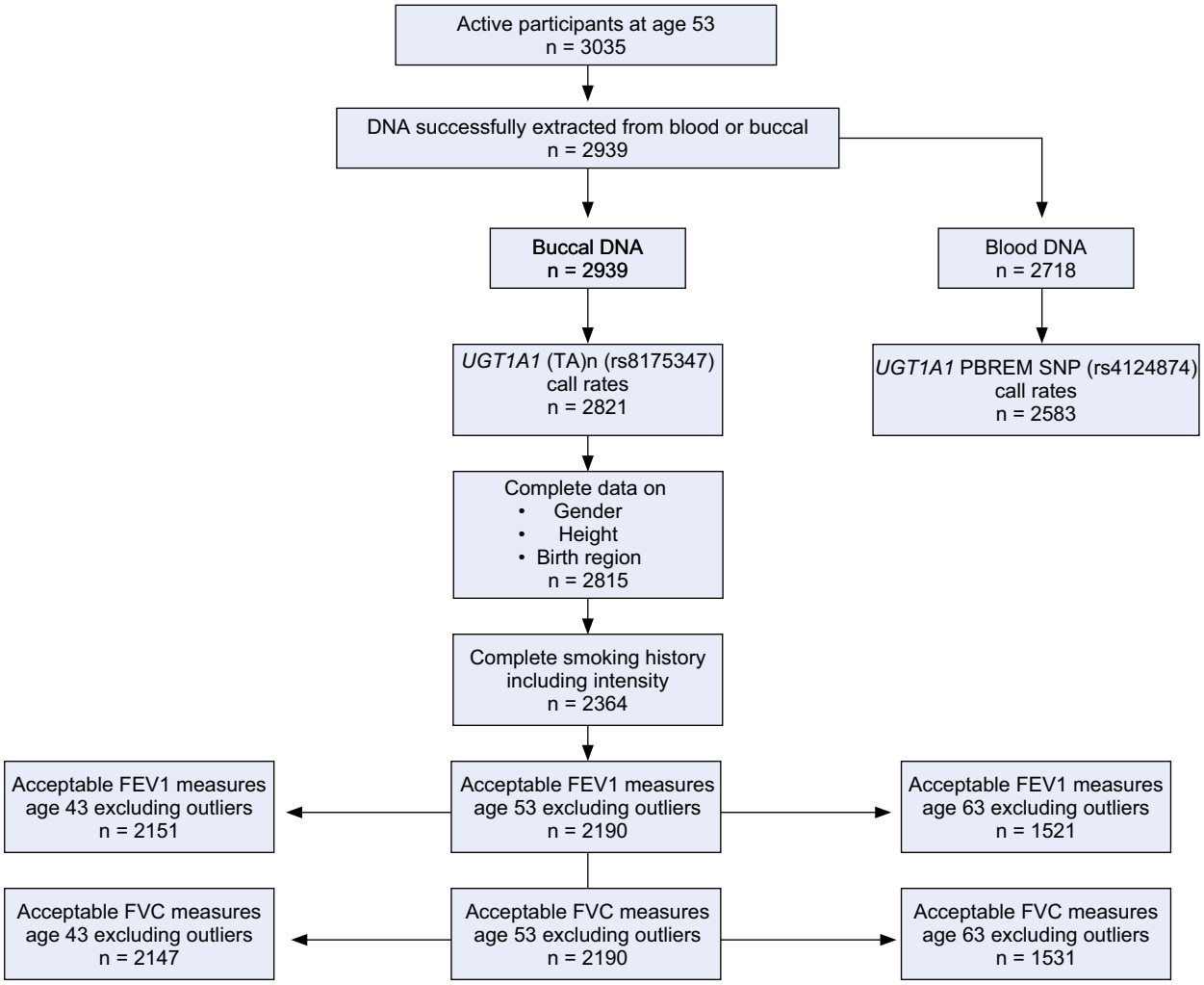

Fig. 1. Flow diagram of cohort participants included in the analyses.

serum bilirubin levels) is 'instrumented' using the genotype (i.e. *UGT1A1* genotypes) to identify and attempt to quantify the genetic component of the variation and thus any causal relationships. Serum bilirubin measures were available for ages 60–64 and these were used in a two-stage instrumental variable analysis to further examine and quantify any causal relationships. Due to lower statistical power and potential for bias from respiratory data missing at this collection point (Supplementary Table 1), the results are included in the Supplementary data section as a supportive analysis.

All analyses were carried out using Stata version 12.1. Probability values of less than 5% (p <0.05) were considered statistically significant assuming a two-sided alternative hypothesis. A higher threshold of less than 10% (p <0.1) was used for detecting potential interaction effects.

Results

In total, complete genotype, respiratory function and covariate data were available for 2190 participants at age 53 when the DNA sample was taken (Fig. 1). The minor allele frequencies for the (TA)n repeat variant (rs8175347) and the PBREM SNP (rs4124874) loci were 0.30 (7 or 8 TA repeats) and 0.43 respectively (Table 1). Consistent with existing literature, serum bilirubin levels measured in a subgroup of participants, remaining in the cohort at ages 60–64, were highest in participants with the *UGT1A1* (TA)n genotype associated with Gilbert's syndrome (Supplementary Patients and methods and Table 1).

Mean FEV1 and FVC was higher for the (TA)n genotype associated with Gilbert's syndrome (group 2; [TA]7/7) with a suggestion that the relationship was stronger in heavy smokers at older ages (Table 2, Fig. 2). Never and ex-smokers at age 36 were pooled because there were no significant differences in respiratory function between these groups. There were no statistically significant univariable associations between *UGT1A1* variation and important predictors of respiratory function including sex, height, smoking status and region of birth.

The (TA)n genotype underlying Gilbert's syndrome (group 2; [TA]7/7) was associated with significantly higher FEV1 (p <0.001) and FVC (p = 0.006) in the overall linear mixed model, including available respiratory data for all ages (Supplementary Table 2). There was no strong indication that respiratory function was lower in the *UGT1A1* high expresser category (group 0) *vs.* the intermediate category (group 1) so we collapsed these into one group for further analyses. The PBREM SNP was not independently related to respiratory function in this mixed linear model after accounting for the (TA)n variation. There was evidence of a genetic interaction with age and smoking status, and thus cross-sectional regression analyses were performed (Table 3).

Table 1. Serum total bilirubin levels by *UGT1A1* (TA)n repeat variant (rs8175347) and the PBREM SNP (rs4124874) measured between ages 60 and 64 in a cohort of British adults.

UGT1A1 rs8175347 (TA)n (%)	Genotype and allele frequencies	Number with bilirubin measures	Mean serum total bilirubin (± 1 SD)
(TA)5/6 - Group 0	7 (0)	4 (0)	7.0 (0.7)
(TA)6/6 - Group 0	1079 (49)	772 (50)	8.0 (3.2)
(TA)5/7 - Group 1	2 (0)	1 (0)	6.6 (0.0)
(TA)6/7 - Group 1	899 (41)	634 (41)	9.6 (4.1)
(TA)6/8 - Group 1	1 (0)	0	
(TA)7/7 - Group 2	202 (9)	140 (9)	18.1 (7.9)
MAF [(TA)7 or (TA)8]	0.30		
UGT1A1 rs4124874 c.-3279T>G (%)*			
TT	711 (33)	518 (34)	7.9 (3.0)
TG	1043 (48)	725 (47)	9.2 (3.9)
GG	416 (19)	295 (19)	13.5 (7.5)
MAF	0.43		

MAF, minor allele frequency; UGT1A1, uridine diphosphate-glucuronosyltransferase 1; SD, standard deviation.
*Numbers genotyped are lower for rs4124874 due to fewer blood samples collected *vs.* buccal samples used to genotype (TA)n.

Table 2. Characteristics of the adult members of the 1946 British birth cohort (MRC NSHD) with DNA samples, adequate respiratory function measures and smoking data recorded at three follow-up periods (ages 43, 53, and 60–64).

Characteristic	Overall n = 2190	Group 0* n = 1088	Group 1* n = 900	Group 2* n = 202
Male gender (%)	1092 (50)	556 (51)	437 (49)	99 (50)
Height, cm	169 (8.8)	169 (8.7)	169 (8.9)	169 (9.1)
FEV1				
Age 43 (n = 2151)	3.02 L (0.70)	3.04 (0.7)	2.97 (0.71)	3.14 (0.69)
Age 53 (n = 2190)	2.81 L (0.70)	2.81 (0.69)	2.78 (0.69)	2.91 (0.69)
Age 60-64 (n = 1521)	2.60 L (0.71)	2.62 (0.71)	2.56 (0.72)	2.69 (0.65)
FVC				
Age 43 (n = 2147)	3.67 L (0.90)	3.7 (0.89)	3.59 (0.89)	3.8 (0.93)
Age 53 (n = 2190)	3.51 L (0.90)	3.52 (0.89)	3.47 (0.91)	3.6 (0.88)
Age 60-64 (n = 1531)	3.29 L (0.90)	3.31 (0.9)	3.25 (0.9)	3.37 (0.88)
Smoking status at age 53 (%)				
Never/ex by age 36	1405 (64)	694 (64)	577 (64)	134 (66)
Light to moderate	558 (26)	291 (27)	222 (25)	45 (22)
Heavy	227 (10)	103 (9)	101 (11)	23 (11)

Total numbers with percentages (in parentheses) are reported for categorical variables and means with ± 1 standard deviation (in parentheses) for continuous variables (height measured in cm and FEV1 and FVC in litres).
FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity.
*Group 0 = (TA)5/6 or (TA)6/6; Group 1 = (TA)5/7, (TA)6/7, (TA)6/8; Group 2 = (TA)7/7 (underlying Gilbert's syndrome).

Overall, the results suggest that the genetic effects on respiratory function become weaker in never/ex-smokers as cohort members age, whereas for heavy smokers the genetic effect seemed to be stronger at older ages. At age 43, the differences in respiratory function between (TA)n groups were fairly similar across smoking categories although there was a suggestion that the magnitude of the difference was greater for heavy smokers (Table 3). At age 53 mean FEV1 was 409 ml higher (p <0.001) in heavy smokers and FVC was 530 ml higher (p <0.001) for those with the genotype underlying Gilbert's syndrome (group 2; [TA]7/7) compared to those without this genotype, whereas the differences for other smoking categories at this age were weaker and non-significant (Table 3). The genetic associations at age 60–64 were also stronger for heavy smokers compared with age 43 but statistical power was lower due to a smaller sample size (Table 3).

Examination of the possible independent association of the PBREM SNP (rs4124874) with respiratory function, after adjustment for the effect of (TA)n showed evidence of some residual effect in heavy smokers at age 53 (p = 0.05 for FEV1 and p = 0.02 for FVC). Mean FEV1 in heterozygotes was 203 ml higher (95%; 0 to 407 ml) and mean FVC was 173 ml higher (95%; −1 to 426 ml) compared with homozygotes for the high expresser allele (data not shown). Including both loci in the regression models for heavy smokers at age 53, increased the explainable variability (R^2) in respiratory function by 9% for FEV1 and 8% for FVC. For comparison, the amount of variability explained by differences in height for this group was 13% for FEV1 and 16% for FVC.

The genotype underling Gilbert's syndrome (group 2; [TA]7/7) was associated with 50% lower odds of overall respiratory disease

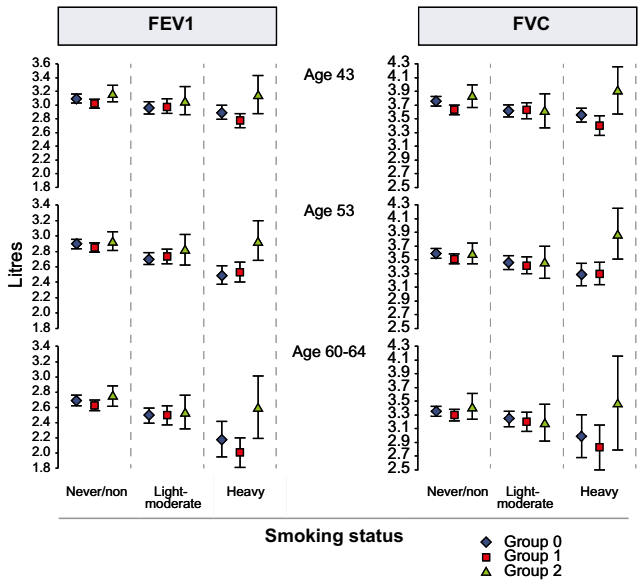

Fig. 2. Mean FEV1 and FVC expressed in litres, with 95% CI by age and smoking status in a cohort of British adults followed since birth. Group 0 = (TA)5/6 or (TA)6/6; Group 1 = (TA)5/7, (TA)6/7, (TA)6/8; Group 2 = (TA)7/7 (underlying Gilbert's syndrome). (This figure appears in colour on the web.)

(Table 4). The effect sizes were similar across COPD, asthma and respiratory drug prescriptions (Table 4). Interactions with smoking status were suggested but were not statistically significant possibly due to lower power for this binary outcome relative to continuous respiratory function.

The instrumental variable analysis including respiratory measurements at all ages and serum bilirubin measures estimated that a 10% increase in serum bilirubin levels was associated with an average 13 ml increase in both FEV (95% CI 4–22 ml; $p = 0.007$) and FVC (95% CI 2–24 ml; $p = 0.02$) (Supplementary Tables 3 and 4). The relationships were strongest in heavy smokers at age 53 with a 10% increase in serum bilirubin level associated with an average 77 ml increase in FEV (95% CI 35–119 ml; $p < 0.001$) and a 79 ml increase in FVC (95% CI 28–130 ml; $p = 0.002$) (Supplementary Tables 3 and 4).

Excluding the small number of participants with rare (TA)5 or (TA)8 alleles had no impact on any of the relationships examined.

Discussion

In the present study, we have shown that participants in the 1946 British birth cohort predisposed to Gilbert's syndrome due to low expression of the enzyme UGT1A1 have higher respiratory function and lower rates of respiratory disease relative to those without this predisposition. The genetic relationships seemed to be strongest amongst heavy smokers after age 43. The similar magnitude of the effect on FEV1 and FVC suggests that *UGT1A1* and unconjugated bilirubin may have a protective role for both obstructive and restrictive respiratory diseases. These results lend support to a causal relationship between raised serum unconjugated bilirubin and protection of the respiratory system against the harmful effects of heavy smoking and show the beneficial effect of Gilbert's syndrome in this context. This protective effect may extend to other environmental sources of oxidants as well as cigarette smoke such as air pollution and to protection from certain types of lung cancer.

To the best of our knowledge, the present study is the first to examine respiratory function and disease in people with the causal variant underlying Gilbert's syndrome and to show an interaction with smoking status. Another group recently reported associations between a SNP, linked to bilirubin levels from genome wide association studies and respiratory function in a Swiss cohort [9]. This study reported interactions with serum bilirubin and smoking status for the full cohort (n = 4195) but lacked power to examine genetic interactions with smoking for the subset with DNA (n = 982). Overall, the genetic effects appear slightly weaker than for our study although differences in age, smoking behaviour and the exclusion of patients with a history of asthma may explain these differences. Genetically reduced haem oxygenase 1 activity, resulting in lower bilirubin levels, has also been associated with poorer respiratory function and a higher risk of respiratory disease with the effect strongest in heavy smokers ($\geqslant$20 cigarettes per day) [32,33]. However, pleiotropy is a possibility for those studies of haem oxygenase 1 where polymorphisms also lower the production of other products of the pathway that may influence respiratory function including carbon monoxide. Genetic variation of *UGT1A1* has been associated with other diseases where oxidative stress is purported to play a pathophysiological role. For example, longitudinal studies with long-term follow-up, including the Framingham Offspring cohort study, have reported 50–80% lower rates of cardiovascular disease and mortality in people with *UGT1A1* (TA)n genotype associated with Gilbert's syndrome or linked loci [34–36]. However, these relationships are not consistently replicated [37].

The strengths of the current study include the long-term follow-up, allowing analysis of respiratory function over a 20-year period, standardized measurement of phenotypes and the use of a strong genetic instrument for bilirubin levels. Limitations include self-reported smoking status, asthma and drug use, and the spirometric definition of COPD without knowledge of symptoms. There may also be some sources of bias. For example, the genetic associations could have been diluted if participants with genetically lower bilirubin levels were more likely to drop out of the cohort before age 53 due to poor health or death before the DNA sample was taken. In the instrumental variable analysis, the stronger causal estimates using genetic instruments *vs.* regression coefficients from using serum bilirubin may suggest that the relationship between serum bilirubin measures and respiratory health are underestimated, due to factors such as regression dilution bias, model misspecification (assuming a linear relationship), reverse causation and negative confounding from drugs and disease processes. Certain antioxidant systems, including haem oxygenase 1 (the first step in bilirubin production), are upregulated in response to disease processes, suggesting that reverse causation is a possibility. Stronger causal estimates from genetic instrumental variable analyses *vs.* other study designs have also been reported with respect to low-density lipoprotein cholesterol and the risk of coronary heart disease [38]. The results of the study are broadly generalizable to the British post-war generation who were mainly of UK/northern European ancestry [17] but the relationships in non-European ethnic groups require further study. Finally, we have used an instrumental variable approach to support causality but this relies on some important assumptions, including absence of confounding by pleiotropy (genes affecting multiple traits), linkage

Table 3. Association between *UGT1A1* rs8175347 (TA)n repeat variants and measures of respiratory function at each age stratified by smoking status in a cohort of British adults.

	FEV1 (ml)			FVC (ml)		
		β coefficients (95%CI)†			β coefficients (95%CI)†	
Variable and regression model	Number	*UGT1A1* (TA)n group 2 *vs.* 0/1*	*p* value	Number	*UGT1A1* (TA)n group 2 *vs.* 0/1*	*p* value
Age 43						
Overall	2151	122 (49 to 194)	0.001	2147	133 (39 to 227)	0.005
Never/ex-smokers	1379	128 (38 to 217)	0.005	1376	159 (45 to 274)	0.006
Light-moderate	462	71 (-92 to 235)	0.392	461	-22 (-233 to 190)	0.841
Heavy	310	205 (14 to 396)	0.035	310	296 (30 to 562)	0.029
Age 53						
Overall	2190	105 (38 to 172)	0.002	2190	92 (8 to 175)	0.032
Never/ex-smokers	1405	79 (-3 to 161)	0.058	1405	68 (-34 to 170)	0.193
Light-moderate	558	63 (-79 to 204)	0.386	558	-35 (-213 to 142)	0.695
Heavy	227	409 (191 to 627)	<0.001	227	530 (262 to 798)	<0.001
Age 60-64						
Overall	1521	98 (10 to 185)	0.029	1531	81 (-26 to 188)	0.138
Never/ex-smokers	1036	87 (-14 to 189)	0.092	1043	86 (-42 to 215)	0.187
Light-moderate	416	111 (-77 to 300)	0.247	417	49 (-162 to 260)	0.649
Heavy	69	538 (46 to 1029)	0.033	71	356 (-339 to 1052)	0.309

FEV1 and FVC are expressed as change in volume measured in ml.
*Group 0 = (TA)5/6 or (TA)6/6; Group 1 = (TA)5/7, (TA)6/7, (TA)6/8; Group 2 = (TA)7/7 (underlying Gilbert's syndrome).
†Regression coefficients for models including sex, height, and birth region.

Table 4. Association between *UGT1A1* rs8175347 (TA)n repeat variants and respiratory disease up to age 60–64 (COPD, asthma, or respiratory drug use) in a cohort of British adults followed since birth in 1946.

	Number of cases (% of cohort)	Adjusted odds ratio (95% CI)† (TA)n group 2 *vs.* 0/1	*p* value
Respiratory disease	478 (21.8)	0.49 (0.32 to 0.74)	<0.001
Spirometrically defined COPD	292 (13.3)	0.49 (0.29 to 0.84)	0.009
Asthma by age 53	211 (9.6)	0.34 (0.17 to 0.70)	0.003
Respiratory drugs	181 (8.3)	0.53 (0.27 to 1.04)	0.064

*Group 0 = (TA)5/6 or (TA)6/6; Group 1 = (TA)5/7, (TA)6/7, (TA)6/8; Group 2 = (TA)7/7 (underlying Gilbert's syndrome).
†Discrete-time proportional odds survival model controlling for sex, height, smoking status, and birth region.

disequilibrium (where due to co-inheritance the variant under study may be a proxy for the causal variant or confounded by a linked variant), and population stratification (spurious associations due to ancestry differences, for example associated with geographic differences in genotype frequencies across study groups) [14]. The 1946 Birth cohort is predominantly white European and we adjusted for possible geographic effects to reduce any potential genetic stratification. The *UGT1A1* gene is located in a gene complex that encodes several other UGT1A isoforms involved in glucuronidation of many other endogenous and exogenous molecules. A region of linkage disequilibrium exists across the complex and encompasses other variants with purported function [39]. Thus, it is plausible that the relationships in the present study reflect substrate effects of a different UGT1A isoform although there are no biologically plausible candidates that we are aware of at present. A further complexity is that antioxidants can increase *UGT1A1* expression as the result of interactions of a signalling pathway, involving the transcription factor Nrf2 via regulatory elements, which overlap with the PBREM region [40].

The amount of variability in respiratory function explained by the *UGT1A1* variation in heavy smokers suggests a role in predicting adult respiratory outcomes and further research is required to investigate this empirically in larger independent cohorts. The results of the present study support but do not prove causality. Randomised controlled studies of the short-term effect of bilirubin infusions or drugs that specifically inhibit UGT1A1 activity on respiratory and endothelial function in humans are perhaps possible, with similar studies carried-out for another serum antioxidant, uric acid, suggesting a physiological benefit [41].

The high frequency of various alleles causing mild hyperbilirubinaemia and Gilbert's syndrome across human populations, including a relatively common non-synonymous SNP of *UGT1A1* in East Asian groups, has led to the speculation that there has been a balancing selection. One theory is that neurotoxicity, induced by very high bilirubin levels in neonates, prevents the alleles causing hyperbilirubinaemia from being more common across populations [42,43]. It should be noted, however, that in certain regions of Equatorial Africa, *UGT1A1* promoter alleles that cause Gilbert's syndrome in Europeans are more common than those associated with normal bilirubin levels [43]. Protection from cigarette and other smoke in the environment of adult modern humans may not have had much impact on reproductive

outcome. Although laboratory research has focused on antioxidant properties of raised bilirubin, protection against respiratory infections, such as pneumonia and bronchitis is also a plausible explanation for the results of the present study [44], and arguably a more compelling hypothesis for the high frequency of alleles causing mild hyperbilirubinaemia rather than protection against chronic late-onset diseases. The potential role of the *UGT1A1* variation and bilirubin levels in the susceptibility to respiratory infections across the life course is an interesting topic for further research.

In summary, a common genetic variation of the *UGT1A1* gene, causing moderately raised bilirubin and Gilbert's syndrome, is associated with significantly higher respiratory function in adults with the strongest effects seen in heavy smokers, and lower rates of respiratory disease. This may reflect antioxidant or other protective properties of serum bilirubin.

Financial support

Dr. Laura Horsfall completed this work as part of a research fellowship funded by the National Institute for Health Research School for Primary Care Research. The National Survey of Health and Development 1946 birth cohort is funded by the Medical Research Council.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Acknowledgments

We thank undergraduate student Seka Lazare who helped with the genotyping.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhep.2014.07.028.

References

[1] Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet 2007;370:765–773.

[2] Juul K, Tybjaerg-Hansen A, Marklund S, Lange P, Nordestgaard BG. Genetically increased antioxidative protection and decreased chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2006;173:858–864.

[3] Nakagami T, Toyomura K, Kinoshita T, Morisawa S. A beneficial role of bile pigments as an endogenous tissue protector: anti-complement effects of biliverdin and conjugated bilirubin. Biochim Biophys Acta 1993;1158:189–193.

[4] Stocker R, Yamamoto Y, McDonagh AF, Glazer AN, Ames BN. Bilirubin is an antioxidant of possible physiological importance. Science 1987;235:1043–1046.

[5] Ryter SW, Morse D, Choi AM. Carbon monoxide and bilirubin: potential therapies for pulmonary/vascular injury and disease. Am J Respir Cell Mol Biol 2007;36:175–182.

[6] Wang HD, Yamaya M, Okinaga S, Jia YX, Kamanaka M, Takahashi H, et al. Bilirubin ameliorates bleomycin-induced pulmonary fibrosis in rats. Am J Respir Crit Care Med 2002;165:406–411.

[7] Horsfall LJ, Rait G, Walters K, Swallow DM, Pereira SP, Nazareth I, et al. Serum bilirubin and risk of respiratory disease and death. JAMA 2011;305:691–697.

[8] Breimer LH, Wannamethee G, Ebrahim S, Shaper AG. Serum bilirubin and risk of ischemic heart disease in middle-aged British men. Clin Chem 1995;41:1504–1508.

[9] Curjuric I, Imboden M, Adam M, Bettschart RW, Gerbase MW, Kunzli N, et al. Serum bilirubin is associated with lung function in a Swiss general population sample. Eur Respir J 2014;43(5):1278–1288.

[10] Kang TW, Kim HJ, Ju H, Kim JH, Jeon YJ, Lee HC, et al. Genome-wide association of serum bilirubin levels in Korean population. Hum Mol Genet 2010;19:3672–3678.

[11] Johnson AD, Kavousi M, Smith AV, Chen MH, Dehghan A, Aspelund T, et al. Genome-wide association meta-analysis for total serum bilirubin levels. Hum Mol Genet 2009;18:2700–2710.

[12] Chen G, Ramos E, Adeyemo A, Shriner D, Zhou J, Doumatey AP, et al. UGT1A1 is a major locus influencing bilirubin levels in African Americans. Eur J Hum Genet 2012;20:463–468.

[13] Bosma PJ, Chowdhury JR, Bakker C, Gantla S, de Boer A, Oostra BA, et al. The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. N Engl J Med 1995;333:1171–1175.

[14] Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32:1–22.

[15] Stafford M, Black S, Shah I, Hardy R, Pierce M, Richards M, et al. Using a birth cohort to study ageing: representativeness and response rates in the National Survey of Health and Development. Eur J Ageing 2013;10:145–157.

[16] Wadsworth M, Kuh D, Richards M, Hardy R. Cohort profile: the 1946 National Birth Cohort (MRC National Survey of Health and Development). Int J Epidemiol 2006;35:49–54.

[17] Wadsworth ME, Butterworth SL, Hardy RJ, Kuh DJ, Richards M, Langenberg C, et al. The life course prospective design: an example of benefits and problems associated with study longevity. Soc Sci Med 2003;57:2193–2205.

[18] Rousseau K, Vinall LE, Butterworth SL, Hardy RJ, Holloway J, Wadsworth ME, et al. MUC7 haplotype analysis: results from a longitudinal birth cohort support protective effect of the MUC7*5 allele on respiratory function. Ann Hum Genet 2006;70:417–427.

[19] Borucki K, Weikert C, Fisher E, Jakubiczka S, Luley C, Westphal S, et al. Haplotypes in the UGT1A1 gene and their role as genetic determinants of bilirubin concentration in healthy German volunteers. Clin Biochem 2009;42:1635–1641.

[20] Lin JP, Cupples LA, Wilson PW, Heard-Costa N, O'Donnell CJ. Evidence for a gene influencing serum bilirubin on chromosome 2q telomere: a genome-wide scan in the Framingham study. Am J Hum Genet 2003;72:1029–1034.

[21] Sugatani J, Yamakawa K, Yoshinari K, Machida T, Takagi H, Mori M, et al. Identification of a defect in the UGT1A1 gene promoter and its association with hyperbilirubinemia. Biochem Biophys Res Commun 2002;292:492–497.

[22] Beutler E, Gelbart T, Demina A. Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? Proc Natl Acad Sci U S A 1998;95:8170–8174.

[23] Chaar V, Keclard L, Diara JP, Leturdu C, Elion J, Krishnamoorthy R, et al. Association of UGT1A1 polymorphism with prevalence and age at onset of cholelithiasis in sickle cell anemia. Haematologica 2005;90:188–199.

[24] Sampietro M, Lupica L, Perrero L, Romano R, Molteni V, Fiorelli G. TATA-box mutant in the promoter of the uridine diphosphate glucuronosyltransferase gene in Italian patients with Gilbert's syndrome. Ital J Gastroenterol Hepatol 1998;30:194–198.

[25] Falaschetti E, Laiho J, Primatesta P, Purdon S. Prediction equations for normal and low lung function from the Health Survey for England. Eur Respir J 2004;23:456–463.

[26] Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.

[27] McClearn GE, Svartengren M, Pedersen NL, Heller DA, Plomin R. Genetic and environmental influences on pulmonary function in aging Swedish twins. J Gerontol 1994;49:264–268.

[28] Gottlieb DJ, Wilk JB, Harmon M, Evans JC, Joost O, Levy D, et al. Heritability of longitudinal change in lung function. The Framingham study. Am J Respir Crit Care Med 2001;164:1655–1659.

[29] Tang W, Bentley AR, Kritchevsky SB, Harris TB, Newman AB, Bauer DC, et al. Genetic variation in antioxidant enzymes, cigarette smoking, and longitudinal change in lung function. Free Radic Biol Med 2013;63:304–312.

[30] Bentley AR, Kritchevsky SB, Harris TB, Newman AB, Bauer DC, Meibohm B, et al. Genetic variation in antioxidant enzymes and lung function. Free Radic Biol Med 2012;52:1577–1583.

[31] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc 1996;91:444–455.
[32] Fredenburgh LE, Perrella MA, Mitsialis SA. The role of heme oxygenase-1 in pulmonary disease. Am J Respir Cell Mol Biol 2007;36:158–165.
[33] Guenegou A, Leynaert B, Benessiano J, Pin I, Demoly P, Neukirch F, et al. Association of lung function decline with the heme oxygenase-1 gene promoter microsatellite polymorphism in a general population sample. Results from the European Community Respiratory Health Survey (ECRHS), France. J Med Genet 2006;43:e43.
[34] Lin JP, O'Donnell CJ, Schwaiger JP, Cupples LA, Lingenhel A, Hunt SC, et al. Association between the UGT1A1*28 allele, bilirubin levels, and coronary heart disease in the Framingham Heart Study. Circulation 2006;114:1476–1481.
[35] Chen YH, Hung SC, Tarng DC. Serum bilirubin links UGT1A1*28 polymorphism and predicts long-term cardiovascular events and mortality in chronic hemodialysis patients. Clin J Am Soc Nephrol 2011;6:567–574.
[36] Cox AJ, Ng MC, Xu J, Langefeld CD, Koch KL, Dawson PA, et al. Association of SNPs in the UGT1A gene cluster with total bilirubin and mortality in the Diabetes Heart Study. Atherosclerosis 2013;229:155–160.
[37] Stender S, Frikke-Schmidt R, Nordestgaard BG, Grande P, Tybjaerg-Hansen A. Genetically elevated bilirubin and risk of ischaemic heart disease: three Mendelian randomization studies and a meta-analysis. J Intern Med 2013;273:59–68.
[38] Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol 2012;60:2631–2639.
[39] Ehmer U, Kalthoff S, Fakundiny B, Pabst B, Freiberg N, Naumann R, et al. Gilbert syndrome redefined: a complex genetic haplotype influences the regulation of glucuronidation. Hepatology 2012;55:1912–1921.
[40] Yueh MF, Tukey RH. Nrf2-Keap1 signaling pathway regulates human UGT1A1 expression in vitro and in transgenic UGT1 mice. J Biol Chem 2007;282:8749–8758.
[41] Waring WS, McKnight JA, Webb DJ, Maxwell SR. Uric acid restores endothelial function in patients with type 1 diabetes and regular smokers. Diabetes 2006;55:3127–3132.
[42] Premawardhena A, Fisher CA, Liu YT, Verma IC, de Silva S, Arambepola M, et al. The global distribution of length polymorphisms of the promoters of the glucuronosyltransferase 1 gene (UGT1A1): hematologic and evolutionary implications. Blood Cells Mol Dis 2003;31:98–101.
[43] Horsfall LJ, Zeitlyn D, Tarekegn A, Bekele E, Thomas MG, Bradman N, et al. Prevalence of clinically relevant UGT1A alleles and haplotypes in African populations. Ann Hum Genet 2011;75:236–246.
[44] Najib F. Defensive role of bilirubinemia in pneumococcal infection. Lancet 1937;229:505–506.