Alpha$_1$-Antitrypsin as a Risk for Infant and Adult Respiratory Outcomes in a National Birth Cohort

Michael E. J. Wadsworth, Lynne E. Vinall, Abigail L. Jones, Rebecca J. Hardy, David B. Whitehouse, Suzanne L. Butterworth, Warren S. Hilder, Jennifer U. Lovegrove, and Dallas M. Swallow

Medical Research Council National Survey of Health and Development, Department of Epidemiology and Public Health, Royal Free Hospital and University College London Medical School, London; and Galton Laboratory, Department of Biology, University College London, London, United Kingdom

Reduced alpha$_1$-antitrypsin (AAT) encoded by the gene *SERPINA1* is a potential risk for pulmonary disease. We investigated *SERPINA1* polymorphism as a risk for infant and adult pulmonary morbidity, and adult respiratory function and its change between 43 and 53 yr. We used data on a British national representative sample (n = 5,362) studied since birth in 1946 to age 53 yr (when n = 3,035), when DNA was first obtained. *SERPINA1* Z and, to a lesser extent, S carriers had an increased risk of infant lower respiratory infection compared with those who were neither S nor Z carriers (Z carriers: odds ratio = 2.32, 95% confidence interval = 1.37–3.92; S but not Z carriers odds ratio = 1.58, 95% confidence interval = 1.10–2.28) after adjustment for environmental, socioeconomic, and developmental factors, and breast-feeding. There was no difference in the adult outcomes at 53 yr according to genotype, nor was there any association of genotype with change in forced expiratory volume at 1 s between 43 and 53 yr. Lower alpha$_1$-antitrypsin, as indicated by carrier status for the Z and S alleles, was a risk for infant lower respiratory infection, but not for adult respiratory outcomes.

A number of genetic, developmental, and environmental factors increase risk of childhood and adult chest disease. One of the most widely studied of the genetic factors is variation in *SERPINA1*, which codes for alpha$_1$-antitrypsin (AAT), a plasma serine protease inhibitor that is synthesized predominantly in the liver and, to a lesser extent, in alveolar macrophages. The main function of AAT in the respiratory tract is to protect the lungs from proteolytic damage caused by neutrophil elastase, a serine protease produced in response to inflammation (1–3). *SERPINA1* is located within a *SERPIN* gene cluster on chromosome 14q32.1 (2). Many *SERPINA1* alleles have been described, with two variants, S and Z, being associated with reduced serum levels of AAT. If individuals with the genotype MM are considered to have normal (i.e., 100%) levels of AAT, the lower levels seen in individuals with genotypes ZZ, SZ, MZ, SS, and MS are ~ 16, 51, 83, 93, and 97%, respectively—the Z allele causing severe AAT deficiency and the S allele causing partial deficiency (4). Serum levels of AAT increase in response to inflammation (2) and insufficient levels of AAT, as can occur in ZZ homozygotes, lead to loss of elasticity and destruction of lung tissue and the development of progressive, irreversible chronic obstructive pulmonary disease (COPD) (2, 5). Risk of COPD is increased by smoking, but particularly so for ZZ homozygotes, because smoking induces an inflammatory response, and polymerization of the AAT Z may exacerbate this response (3).

In addition, respiratory function, as measured by forced expiratory volume (FEV), which normally declines with age, declines more rapidly in ZZ homozygotes, and the effect is worsened by cigarette smoking (6), although reduced lung function was not found to be associated with *SERPINA1* homozygosity in young adults (7). Signs of asthma are more common in AAT-deficient adults, particularly in homozygotes (7, 8). However, although Z carrier smokers are reported to be at greater risk of decline in FEV at 1 s (FEV$_1$) than non–Z carrier smokers (6), other studies have not found lung function to be associated with *SERPINA1* heterozygosity in adulthood (9). Similarly, some studies have shown the Z allele to be overrepresented in patients with COPD (10–14), and Seersholm and colleagues found that individuals with the MZ genotype who were first-degree relatives of ZZ patients with COPD were at increased risk of hospital admission for COPD (15). In contrast, other studies found that MZ individuals carried no greater risk for developing lung disease than nondeficient individuals when corrected for age, race, sex, and smoking history (16–18). The S allele was not found to be significantly associated with predisposition to COPD or with reduced FEV$_1$ when compared with normal individuals (11, 12, 19).

It has also been reported that the A allele of a G/A single nucleotide polymorphism, situated within an enhancer element 1,237 nucleotides 3′ of the *SERPINA1* gene, is significantly more frequent in patients with COPD than in healthy control subjects (20, 21), although these findings have also been conflicting, with no association found in other studies (10, 22). The presence of this allele is thought to be associated with reduced levels of binding of a transcription factor and an impaired inflammatory response (1, 2).

Developmental and environmental factors are also associated with risk of COPD and poor respiratory function in adulthood. Epidemiologic studies show that low birth weight and childhood pulmonary disease are important factors in the progression to these adult outcomes (23–25). Poor socioeconomic environment in childhood, high exposure to parental smoking and atmospheric pollution, and a short or nonexistent period of exclusive breastfeeding increase the risk associated with developmental sources (24, 25).

Availability of DNA for the first time in a national birth cohort study allows us to assess the role of genetic polymorphism in relation to developmental and environmental factors and outcome measures of infant pulmonary disease and adult respiratory health in a prospectively studied, nationally representative population, selected only by date of birth.

Materials and Methods

The Sample

The Medical Research Council National Survey of Health and Development is a birth cohort study of a sample (n = 5,362) of all births that

(*Received in original form January 27, 2004 and in revised form July 19, 2004*)

Address correspondence to: Michael E. J. Wadsworth, Medical Research Council National Survey of Health and Development, Department of Epidemiology and Public Health, Royal Free Hospital and University College London Medical School, 1-19 Torrington Place, London WC1E 6BT, UK. E-mail: m.wadsworth@ucl.ac.uk

Abbreviations: alpha$_1$-antitrypsin, AAT; chronic obstructive pulmonary disease, COPD; forced expiratory volume at 1 s, FEV$_1$; forced vital capacity, FVC; lower respiratory infection, LRI; Global Initiative for Chronic Obstructive Lung Disease, GOLD.

Am. J. Respir. Cell Mol. Biol. Vol. 31, pp. 559–564, 2004
Originally Published in Press as DOI: 10.1165/rcmb.2004-0027OC on July 22, 2004
Internet address: www.atsjournals.org

TABLE 1. Data collections and sample size

Data Collection	Age (*yr*)	Year	Total Sample Providing Data (*n*)	Available Sample Providing Data[‖] (%)
First	Birth*	1946	5,362	100
Second	2[†]	1948	4,689	94
Third	4[‡]	1950	4,700	96
Twenty-first	43[§]	1989	3,262	87
Twenty-second	53[¶]	1999	3,035	83

* Birth weight, birth order, region of birth, and sex abstracted from medical records.

† Height and weight measured at home or clinics by community nurses, who also asked questions on infant feeding, crowding, and LRI; atmospheric pollution data was sourced from official statistics.

‡ Fathers socioeconomic circumstances (i.e., social class).

§ FEV1, FVC measured by research nurses at home visits.

¶ FEV1, FVC and height measured by research nurses; blood and buccal samples were taken to prepare a DNA source, and questions on wheezing and own and parents smoking were asked by research nurses at home visits.

‖ At each data collection, percentage was calculated by the numerator being the total sample providing data and the denominator being the number of sample members currently available (i.e., excluding the deceased, those who refused further contact, and those resident outside England, Wales, or Scotland).

occurred in England, Wales, and Scotland in 1 wk in March 1946 (26). The sample includes all single, legitimate births whose fathers were in nonmanual or agricultural occupations and a randomly selected one in four of all other single, legitimate births (26). The sampling predates the major postwar immigration into Britain, and there are no non-European sample members. The sample successfully contacted in adulthood is representative of the national population of a similar age (26). Information on health, development, education, and occupation has been collected 22 times (Table 1). Research nurses visited homes to measure a range of health outcomes and to collect data on socioeconomic circumstances of subjects in adulthood. At the most recent collection, when subjects were 53-yr-old, nurses collected information from 3,035 cohort members, 83% of the 3,673 individuals who were still alive, resident in Britain, and who had not previously refused further contact ($n = 640$). At age 53 yr, we did not attempt to contact 31% (1,689) of the sample selected at birth (5,362). Either because they had, by that age, died (9%), were residing abroad (10%), or already refused all further contact (12%).

In this analysis the sample was, for most purposes, restricted to those who responded to the data collection at age 53 yr, because that was when a source of DNA was first collected. The aspect of analysis concerned with mortality used the whole cohort selected at birth.

The Measures

Childhood health, developmental, and environmental measures. Infant lower respiratory infection (LRI) experienced by age 2 yr was reported by mothers at interviews with health visitors when the child was 2 yr old and coded here as none/any. Reports of hospital admissions were checked with hospital records. This variable includes all reports of bronchitis, bronchiectasis, pneumonia of any type, or bronchopneumonia during the first 2 yr of life; the constituent parts of this cannot be analyzed separately because of the wording of the question "Has this baby ever had a lower respiratory infection, e.g., bronchitis, broncopneumonia or pneumonia?"

Information on prenatal growth was provided by birth weight (g) recorded by midwives and health visitors from records when the child was aged 8 wk (information on state of maturity is not available). Postnatal growth was assessed in terms of weight (g) and height (cm) measured by health visitors when the child was aged 2 yr.

Infant feeding history was collected by health visitors at home visits when the child was aged 2 yr and coded here as breast fed for 1–4 mo, $\geqslant$ 5 mo, or bottle fed.

Environmental data included information on parental smoking (asked in retrospect at age 53 yr: "Did either of your parents smoke cigarettes, cigars or pipes when you lived with them as a child?", and was coded as mother smoked yes/no, father smoked yes/no), and on atmospheric pollution from coal burning, which was taken from national official sources (23), summed for the years 1946 to 1948 (ages 0 to 2 yr), and used as a continuous variable. Socioeconomic and family data were crowding (0.5 persons per room; 1 person per room; $>$ 1 person per room), birth order (1, 2, $\geqslant$ 3), father's occupational social class when the child was aged 2 yr (nonmanual/manual), and parental highest educational attainment (neither has qualifications, either or both have some post school but below tertiary [university] level qualification, either or both have tertiary qualifications). Geographical data were place of birth coded into 10 census regions.

Adult health, environment, and survival. Research nurses, trained by the research team, measured FEV_1 and forced vital capacity (FVC) at home visits when subjects were 43- and 53-yr old, using the Micromedical (UK) turbine electronic spirometer. Two measurements were taken, and the maximum readings are presented here. The difference in FEV_1 at 43 and 53 yr is the decline in FEV_1 over that 10-yr age period. When subjects were age 53 yr, the nurses also asked "Does your chest ever sound wheezy or whistling?" (yes/no) and, if "yes," "Do you get this most days or nights?" (yes/no). At the same visit, the nurses measured height using a portable CMS (UK) stadiometer, with the subjects head in the Frankfort plane. Subject smoking history was compiled from data collected at 53 yr and coded as never smoked, ever regularly smoked $>$ 1 cigarette/d but not now, now smoking 1–10 cigarettes/d, 11–20/d, or $>$ 20/d. Information on death was obtained from the National Health Service Register, on which all sample members are flagged, and causes and date confirmed by death certificate.

The genetic measures. At the same visit, the nurses also collected blood and buccal samples from 91 and 96% of respondents, respectively. *SERPINA1* phenotypes were determined by isoelectric focusing of AAT protein present in ethylenediamine tetraacetic acid plasma, as described by Whitehouse and colleagues (27), with the following exceptions. After treatment with DTT and IAA the samples were not further treated. Solutions used for contact between the gel and the electrodes were 0.1 M sodium hydroxide for the cathode and 0.04 M glutamic acid for the anode. Samples (5 μl) were applied to the gel using a multichannel pipette and an Immobiline sample application strip (Amersham Pharmacia Biotech, Buckinghamshire, UK), which was left in place throughout the run. This allowed us to distinguish four different classes of M allele as well as three other rarer alleles and S and Z. However, for most purposes, these protein alleles were classified as either S, Z, or M (to include all M and other rare alleles). Buccal DNA was prepared using DNA extraction solution kindly provided by Ian Craig (Institute of Psychiatry, University of London, UK). Blood DNA was prepared using the Puregene DNA Isolation Kit (Flowgen, Leicestershire, UK) according to the instructions of the manufacturer. The G1237A polymorphism was analyzed by polymerase chain reaction followed by restriction digestion with *Taq*I, as previously described by Sandford and colleagues. (22).

Statistical Analysis

Infant LRI by age 2 yr was the early-life outcome. Five adult outcomes were examined, each using measurements made at age 53 yr. The first outcome was the ratio FEV_1/FVC. The second outcome was the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria for COPD (28), with *moderate* ($FEV_1/FVC < 70$, $30 \leqslant FEV_1 < 80\%$ predicted) combined with *severe* ($FEV_1/FVC < 70\%$, $FEV_1 < 30\%$ predicted or $FEV_1 < 50\%$ predicted plus respiratory failure or clinical signs of right heart failure) and compared with all others. Prediction equations were those from the Health Survey for England for men and women aged $>$ 25 yr (29). The third adult outcome was reported wheeze at 53 yr, the fourth was mean FEV_1, and the fifth was decline in FEV_1 as measured by the difference in FEV_1 between 43 and 53 yr. *SERPINA1* was categorized in three groups, as neither S nor Z carriers (noncarriers), S and not Z carriers, including all six SS homozygotes (S carriers), and Z carriers. ZZ homozygotes and SZ heterozygotes were too few ($n = 2$ and $n = 4$, respectively) to analyze separately, and were included in the Z carriers.

The population size for each protein allele of *SERPINA1* and for G1237A is given in Table 2. The first analysis examined whether the genotype distributions of the G1237A polymorphism and the M, S, and Z protein alleles deviated from the Hardy-Weinberg equilibrium. Then we examined the allelic association of G1237A with the protein alleles of *SERPINA1*, using the expectation maximization (EM) algorithm from the Arlequin program (30).

Next, the association of *SERPINA1* with the outcomes was exam-

TABLE 2. Association of G1237A with the protein alleles of *SERPINA1* and distribution of numbers of each protein allele

	SERPINA1 Genotypes			
	G1237A			
Protein Alleles	GG	GA	AA	Total No. (%)
MM	1,997	321	18	2,336 (85.9)
MS	238	27	0	265 (9.7)
SS	6	0	0	6 (0.2)
MZ	99	8	0	107 (3.9)
SZ	4	0	0	4 (0.2)
ZZ	1	1	0	2 (0.1)
Totals	2,345	357	18	2,720 (100.0)

Numbers of respondents shown. All nondeficient alleles classed as M. Total population tested = 2,720. The observed numbers did not deviate from those expected under the Hardy-Weinberg equilibrium (P = 0.69), G1237A (P = 0.29).

ined using χ^2 and t tests. Then the association of infant LRI with the geographic, social, developmental, and environmental factors was examined, and the geographic distribution of carrier status was compared with that of infant LRI using univariate logistic regression and χ^2 tests. Logistic regression analysis was then used to test whether the association of carrier status with infant LRI could be accounted for by the developmental, social, and environmental factors. First, an unadjusted model was used; then separate models, adjusting for social, developmental, nutritional, and environmental factors; and finally a model adjusting for all factors. Interactions of *SERPINA1* with atmospheric pollution and with parental smoking were tested, because of their inflammatory action, and because other studies (6) report greater FEV_1 decline in *SERPINA1* Z carrier smokers than in *SERPINA1* Z carrier nonsmokers.

The associations of *SERPINA1* with FEV_1, the GOLD criteria, FEV1/FVC < 70%, and reported wheeze, all at 53 yr, were examined using logistic and normal regression analysis as appropriate, adjusting for height, sex, and smoking history. These analyses were repeated with the G1237A polymorphism.

Finally, the difference between FEV_1 at 43 and 53 yr was examined in relation to *SERPINA1* carrier status, and a conditional regression model was generated with FEV_1 at 53 yr as the outcome, adjusting for FEV_1 at 43 yr, sex, adult height, and smoking history. The interaction of FEV_1 decline with smoking history was tested.

Because loss to follow-up is an inevitable problem in long-term longitudinal studies, we looked at how that might have affected the results. First, we used a Cox proportional hazard model to see whether loss through death in the population of subjects who had had infant LRI differed from that in the population of those who did not have that illness. Second, we investigated whether the population for which we have *SERPINA1* data at 53 yr (90% of those providing information at that age) differed from others not seen at that age in terms of known respiratory risk factors and experience of infant LRI.

The level of statistical significance was taken as $P = 0.05$ throughout. Analyses were performed using SPSS, Inc. (Chicago, IL). The n values differ between analyses because data are taken from several data collections with missing values in each, with consequent reduction in n in analyses using data from > 1 collection.

Results

The distribution of demographic characteristics of the sample is given in the Appendix Table A1.

Allelic Distribution in the Sample

Table 2 shows the distribution of the different protein phenotypes (interpreted as genotypes) and also the G1237A genotypes in the whole cohort. Four percent of the sample had the MZ genotype, and 0.3% were SZ or ZZ. Two analyses of fit to Hardy-Weinberg equilibrium were performed. These were for the three protein alleles (M, S, and Z), which were considered together as one gene locus for this test because they were tested by a single method isoelectric focussing (IEF), and for the G1237A polymorphism that was tested by another method (as a DNA single nucleotide polymorphism), and for which two alleles were tested. These numbers do not deviate from those expected under Hardy-Weinberg equilibrium for any polymorphism (M, S, or Z protein alleles, P = 0.69; G1237A, P = 0.29).

TABLE 3. Association of G1237A with the protein alleles of *SERPINA1*

	Haplotype Frequencies			
	ML		No Association	
	G	A	G	A
M	0.856	0.071	0.860	0.067
S/Z	0.071	0.002	0.068	0.005

Definition of abbreviations: EM, expectation maximization; ML, maximum likelihood haplotype frequencies.

ML generated using EM and frequencies in the absence of association. Alleles classified as nondeficient, M (M and rare alleles), or deficient, S/Z; n = 5,440 alleles; the polymorphisms are significantly associated (P = 0.00184).

Allelic Association

The G1237A A allele was associated more frequently than expected with M (P = 0.002) (Table 3), and more particularly with M1 (P < 0.001; data not shown).

Infant LRI and *SERPINA1*

SERPINA1 Z and S carriers each had a raised risk of infant LRI compared with noncarriers (Table 4). The risk was highest in Z carriers, lowest in noncarriers, and intermediate in S carriers. The G1237A polymorphism was not associated with infant LRI (data not shown).

The Association of *SERPINA1* and Environmental Factors with Infant LRI

Infant LRI was associated with all the social factors ($P \leq 0.01$), but was not associated with sex. The developmental factors were not associated with infant LRI, but breast-feeding was a significant protective factor (P = 0.02). Atmospheric pollution and parental smoking were risks ($P \leq 0.001$ and $P \leq 0.04$, respectively). Table 5 shows that in all but two regions, risk of infant LRI was higher in S and Z carriers. However, this was a statistically significant risk in only two regions, as the number of Z carriers was too small for this analysis, as the wide confidence intervals show. The geographic distribution of carrier status and of infant LRI is given in Appendix Table A2.

In an unadjusted logistic regression model, *SERPINA1* was associated with infant LRI (Table 6), and that association remained significant in models adjusting for each group of risk factors and for all factors (Table 6). The unadjusted analysis shown in Table 6 was repeated (not shown) excluding the ZZ homozygotes and SZ heterozygotes ($n = 6$) in order to determine whether the effect of those high-risk genotypes accounted for the findings, but they did not; infant LRI remained strongly associated with the three groups (noncarriers odds ratio [OR] = 1.00, S carriers OR = 1.35 [95% confidence interval {CI} = 0.96–1.88], Z carriers OR=2.34 [95% CI=1.46–3.77] P = 0.001). Rerunning each of the models shown in Table 6, allowing inclusion of all possible sample members in each (not shown), yielded similar results to those presented in Table 6.

Tests for interactions of *SERPINA1* with maternal and paternal smoking and with atmospheric pollution were not statistically significant (P = 0.99, P = 0.98, and P = 0.22, respectively, data not shown).

TABLE 4. *SERPINA1* in relation to the outcome variables

	Protein Alleles				
	Noncarriers % (*n*)*	S Carriers % (*n*)*	Z Carriers % (*n*)*	Total	Significance†
With infant LRI (0–2 yr)	23.5 (2,179)	30.2 (245)	39.0 (105)	2,529	$P = < 0.001$
FEV_1/FVC < 70%	9.7 (2,265)	10.3 (261)	8.0 (112)	2,638	$P = 0.790$
With GOLD criteria moderate or severe COPD:					
Men	9.2 (1,127)	9.8 (133)	6.3 (48)	1,308	$P = 0.760$
Women	5.4 (1,130)	6.3 (128)	3.1 (64)	1,322	$P = 0.660$
With wheeze day or night at 53 yr	9.0 (2,344)	5.5 (272)	7.1 (113)	2,729	$P = 0.130$
FEV_1 at 53 yr, mean (SD, *n*)	2.79 (0.70, 2267)	2.82 (0.73, 260)	2.83 (0.67, 112)	2,639	$P = 0.360$

Definition of abbreviations: FEV_1, forced expiratory volume at 1 s; GOLD, Global Initiative for Chronic Obstructive Lung Disease; LRI, lower respiratory infection.
* Percent of the given population in each cell with the specified condition.
† χ^2 test, except in the comparisons of mean FEV_1 where *t* test was used.

SERPINA1 and Adult Respiratory Health

By age 53 yr, there were no associations of *SERPINA1* with the adult respiratory outcomes (Table 4), nor were there after adjusting for sex, height, and smoking (not shown), nor were there any associations with the G1237A polymorphism (not shown).

Although mean decline in FEV_1 between ages 43 and 53 yr was greatest in the Z carrier group, differences in mean decline across the three categories of *SERPINA1* carrier status were not significant ($P = 0.49$, data not shown), nor was a regression model of carrier status in relation to conditional change in FEV_1 ($P = 0.65$). There was no interaction between FEV_1 decline and smoking history ($P = 0.89$).

Mortality in Relation to Infant LRI

In the total sample of subjects for whom information on infant LRI was available ($n = 4,679$), the rate of death between ages 2 and 53 yr was significantly greater in those with that illness (hazard ratio = 1.53, CI = 1.21–1.95, $P = 0.001$).

Missing Sample Members

Comparison of those subjects from which *SERPINA1* data were collected at 53 yr with all other sample members showed that the population not included in our analysis had significantly lower mean FEV_1 at 43 yr ($P = 0.006$), a greater likelihood of smoking at that age ($P < 0.001$), and higher exposure to atmospheric pollution at 0–2 yr ($P = 0.05$), but did not differ in prevalence of infant LRI ($P = 0.65$) or in mean height at 2 yr ($P = 0.11$).

TABLE 5. Logistic regression analysis showing risk of infant lower respiratory infection by carrier status within region of birth

Region	*n*	S carriers OR (*95% CI*)	Z carriers OR (*95% CI*)	*P* Value for Wald Test
S. England	132	*	*	*
N. England	199	2.06 (0.80–5.29)	3.78 (1.02–14.00)	0.06
E. and W. Yorkshire	177	0.91 (0.32–2.64)	3.29 (0.45–24.19)	0.49
N.W. England	263	1.46 (0.60–3.56)	1.26 (0.29–5.43)	0.68
N. Midlands	209	1.62 (0.60–4.42)	1.70 (0.46–6.27)	0.49
Midlands	197	1.14 (0.41–3.19)	2.28 (0.14–37.09)	
E. England	187	3.91 (1.42–10.73)	2.69 (0.84–8.56)	0.01
London and S.E.	592	1.23 (0.69–2.21)	1.64 (0.68–3.97)	0.45
S.W. England	143	0.90 (0.24–3.43)	1.96 (0.46–8.40)	0.64
Scotland	275	1.76 (0.47–6.58)	5.86 (1.69–20.35)	0.02
Wales	142	0.99 (0.30–3.30)	3.96 (0.84–18.69)	0.21

Definition of abbreviations: CI, confidence interval; OR, odds ratio.
* Cannot be calculated because no Z carriers had infant lower respiratory infection.
Noncarrier reference categories were all 1.00.

Discussion

To our knowledge, this is the first study to examine the relationship of *SERPINA1* with infant and adult respiratory outcomes in the same sample. We have the advantage of measurements and data collection on risk exposures made throughout the life of sample members in this longitudinal study (birth to 53 yr), with the exception of data on parental smoking, which were collected in retrospect. The DNA resource from which *SERPINA1* information was derived was collected at age 53 yr.

We show that *SERPINA1* played a significant and independent role in the risk of infant LRI, but there was no interaction with atmospheric pollution, as might be expected from the find-

TABLE 6. Unadjusted and adjusted logistic regression models of *SERPINA1* in relation to infant (age 0–2 yr) lower respiratory infection ($n = 1,653$)

	OR (*95% CI*)	*P* Value for Wald Test
Unadjusted		
SERPINA1		
Noncarriers	1.00	< 0.001
S carriers	1.63 (1.15–2.31)	
Z carriers	2.13 (1.30–3.44)	
Adjusted for social factors*		
SERPINA1		
Noncarriers	1.00	< 0.001
S carriers	1.57 (1.09–2.25)	
Z carriers	2.32 (1.38–3.90)	
Adjusted for developmental measures and breast feeding†		
SERPINA1		
Noncarriers	1.00	
S carriers	1.66 (1.16–2.36)	< 0.01
Z carriers	2.13 (1.29–3.51)	
Adjusted for environmental factors‡		
SERPINA1		
Noncarriers	1.00	< 0.001
S carriers	1.61 (1.13–2.29)	
Z carriers	2.17 (1.32–3.57)	
Adjusted for G1237A polymorphism		
SERPINA1		< 0.01
Noncarriers	1.00	
S carriers	1.63 (1.15–2.31)	
Z carriers	2.14 (1.30–3.51)	
Fully adjusted		
SERPINA1		
Noncarriers	1.00	< 0.001
S carriers	1.58 (1.10–2.28)	
Z carriers	2.32 (1.37–3.92)	

Definition of abbreviations: CI, confidence interval; OR, odds ratio.
* Social factors: sex, father's social class, crowding at 2 yr, region of birth, birth order, and parental educational attainment.
† Developmental and nutritional factors: birth weight, height at 2 yr, weight at 2 yr, and infant feeding.
‡ Environmental factors: atmospheric pollution 0–2 yr, father's smoking, and mother's smoking.

ings of von Ehrenstein and colleagues (31). The results suggest that the level of AAT may be critical during this developmental period (1).

Although we found no evidence of interactions between parental smoking and *SERPINA1* in relation to infant LRI, this may not be a reliable result, as parental smoking was recollected. The absence of an interaction with atmospheric pollution may be the result of the small number of Z carriers.

The absence of association of *SERPINA1* with adult respiratory outcomes was similar to the findings of others, who also failed to show an association with heterozygosity for *SERPINA1* S and Z (9, 16–18). We also failed to find an association with reported wheeze, which others have found in those severely deficient in AAT (7, 8). Our findings may reflect the nature of our cohort in which, although relatively large, only 6 individuals would be expected to be severely deficient, and the fact that, unlike other large studies, our sample was not taken from a population of patients. We have no information on AAT deficiency in relatives, which Seersholm and colleagues (15) showed to be a risk. Our study is thus similar to the sample used by Silva and colleagues (9), who also reported no association, but differs from that of several hospital-based studies, in which associations were found (10–15, 19). It may be that in this sample at later ages, as the risk of COPD and the slope of functional decline each increase, their association with *SERPINA1* will change (10).

Those for whom we have no DNA sample at 53 yr, but who had provided data 10 yr earlier, had lower mean FEV_1 at that age ($P = 0.01$), suggesting that they would also have had a lower FEV_1 at 53 yr. This, together with the significantly higher exposure to atmospheric pollution in infancy among those not providing data at 53 yr ($P = 0.05$), and a greater extent of contact loss among those who were smokers at 43 yr ($P < 0.001$), indicates that the effects of smoking and infant exposure to atmospheric pollution may have been underestimated in our analyses. The significantly elevated risk of death before age 53 yr (when the source of DNA was first collected) among those who had infant LRI indicates that the observed associations between *SERPINA1* alleles and the disease outcomes in infancy and adulthood are likely to be weaker than the true effect of *SERPINA1* on respiratory function. However, death rates by any cause in this population are representative of those in the national population of a similar age (26).

Conclusions

SERPINA1 was shown to be a risk for infant LRI in this community sample, independently of environmental factors. However, unlike infant LRI, *SERPINA1* was shown not to be a significant risk for adult respiratory outcomes by age 53 yr. There was no significant association of the G1237A polymorphism with either infant LRI or adult respiratory outcomes.

Conflict of Interest Statement: M.E.J.W. has no declared conflicts of interest; L.E.V. has no declared conflicts of interest; A.L.J. has no declared conflicts of interest; R.J.H. has no declared conflicts of interest; D.B.W. has no declared conflicts of interest; S.L.B. has no declared conflicts of interest; W.S.H. has no declared conflicts of interest; J.U.L. has no declared conflicts of interest; and D.M.S. has no declared conflicts of interest.

Acknowledgments: The authors are grateful to Prof. Sue Povey of the Department of Biology at University College London for comments and encouragement; to the Medical Research Council for funding the MRC National Survey of Health and Development, and the collection of data and material for making DNA; to Prof. Ian Craig and Mr. Bernard Freeman at the Molecular Genetics Group, Institute of Psychiatry, London for their advice and help with buccal DNA preparation; and to the National Centre for Social Research, whose research nurses collected the data and biological samples.

References

1. Barnes, P. J. 1999. Genetics and pulmonary medicine: molecular genetics of chronic obstructive pulmonary disease. *Thorax* 54:245–252.
2. Kalsheker, N., S. Morley, and K. Morgan. 2002. Gene regulation of the serine proteinase inhibitors alpha1-antitrypsin and alpha1-antichymotrypsin. *Biochem. Soc. Trans.* 30:93–98.
3. Lomas, D. A., and R. Mahadeva. 2002. Alpha 1-antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy. *J. Clin. Invest.* 110:1585–1590.
4. Brantly, M. L., J. T. Wittes, C. F. Vogelmeier, R. C. Hubbard, G. A. Fells, and R. G. Crystal. 1991. Use of a highly purified alpha1-antitrypsin standard to establish ranges for the common, normal and deficient alpha1-antitrypsin phenotypes. *Chest* 100:703–708.
5. Dabbagh, K., G. J. Laurent, A. Shock, P. Leoni, J. Papakrivopoulou, and R. C. Chambers. 2001. Alpha-1–antitrypsin stimulates fibroblast proliferation and procollagen production and activates classical MAP kinase signalling pathways. *J. Cell. Physiol.* 186:73–81.
6. Sandford, A. J., T. Chagani, T. D. Weir, J. E. Connett, N. R. Anthonisen, and P. D. Pare. 2001. Susceptibility genes for rapid decline of lung function in the Lung Health Study. *Am J Respir Crit Care Med* 163:469–473.
7. Piitulainen, E., and T. Sveger. 2002. Respiratory symptoms and lung function in young adults with severe alpha (1)-antitrypsin deficiency (PiZZ). *Thorax* 7:705–708.
8. Eden, E., J. Hammel, F. N. Rouhani, M. L. Brantly, A. F. Barker, A. S. Buist, R. J. Fallat, J. K. Stoller, R. G. Crystal, and G. M. Turino. 2003. Asthma features in severe alpha1-antitrypsin deficiency. *Chest* 123:765–771.
9. Silva, G. E., D. L. Sherrill, S. Guerra, and R. A. Barbee. 2003. A longitudinal study of alpha1-antitrypsin phenotype and decline in FEV1 in a community population. *Chest* 123:1435–1440.
10. Dahl, M., A. Tybjaeg-Hansen, P. Lange, J. Vetsbo, and B. Nordestgaard. 2002. Change in lung function and morbidity from chronic obstructive pulmonary disease in alpha1-antitrypsin MZ heterozygotes: a longitudinal study in the general population. *Ann. Intern. Med.* 136:270–279.
11. Alvarez-Granda, L., M. J. Cabero-Perez, A. Bustamente-Ruiz, D. Gonzalez-Lamuno, M. Delgado-Rodriguez, and M. Garcia-Fuentes. 1997. PI SZ phenotype in chronic obstructive pulmonary disease. *Thorax* 52:659–661.
12. Lieberman, J., B. Winter, and A. Sastre. 1986. Alpha 1-antitrypsin Pi-types in 965 COPD patients. *Chest* 89:370–373.
13. Bartmann, K., M. Fooke-Achterrath, G. Koch, I. Nagy, I. Schutz, E. Weis, and M. Zierski. 1985. Heterozygosity in the Pi-system as a pathogenetic cofactor in chronic obstructive pulmonary disease (COPD). *Eur. J. Respir. Dis.* 66:284–296.
14. Cox, D. W., V. H. Hoeppner, and H. Levison. 1976. Protease inhibitors in patients with chronic obstructive pulmonary disease: the alpha-antitrypsin heterozygote controversy. *Am. Rev. Respir. Dis.* 113:601–606.
15. Seersholm, N., J. T. Wilcke, A. Kok-Jensen, and A. Dirksen. 2000. Risk of hospital admission for obstructive pulmonary disease in α1-antityrpsin heterozygotes of phenotype PiMZ. *Am J Respir Crit Care Med* 161:81–84.
16. Bruce, R. M., B. H. Cohen, E. L. Diamond, R. J. Fallat, R. J. Knudson, M. D. Lebowtiz, C. Mittman, C. D. Patterson, and M. S. Tockman. 1984. Collaborative study to assess risk of lung disease in Pi MZ phenotype subjects. *Am. Rev. Respir. Dis.* 130:386–390.
17. Morse, J. O., M. D. Lebowitz, R. J. Knudson, and B. Burrows. 1977. Relation of protease inhibitor phenotypes to obstructive lung diseases in a community. *N. Engl. J. Med.* 296:1190–1194.
18. Eriksson, S., S. E. Lindell, and R. Wiberg. 1985. Effects of smoking and intermediate alpha 1-antitrypsin deficiency (PiMZ) on lung function. *Eur. J. Respir. Dis.* 67:279–285.
19. Sandford, A. J., T. D. Weir, J. J. Spinelli, and P. D. Pare. 1999. Z and S mutations of the alpha1-antitrypsin gene and the risk of chronic obstructive pulmonary disease. *Am. J. Respir. Cell Mol. Biol.* 20:287–291.
20. Kalsheker, N. A., G. L. Watkins, S. Hill, K. Morgan, R. A. Stockley, and R. B. Fick. 1990. Independent mutations in the flanking sequence of the alpha-1–antitrypsin gene are associated with chronic obstructive airways disease. *Dis. Markers* 8:151–157.
21. Poller, W., C. Meisen, and K. Olek. 1990. DNA polymorphisms of the alpha 1-antitrypsin gene region in patients with chronic obstructive pulmonary disease. *Eur. J. Clin. Invest.* 20:1–7.
22. Sandford, A. J., J. J. Spinelli, T. D. Weir, and P. D. Pare. 1997. Mutation in the 3′ region of the alpha-1–antitrypsin gene and chronic obstructive pulmonary disease. *J. Med. Genet.* 34:874–875.
23. Mann, S. L., M. E. J. Wadsworth, and J. R. T. Colley. 1992. Accumulation of factors influencing respiratory illness in members of a national birth cohort and their offspring. *J. Epidemiol. Community Health* 46:286–292.
24. Shaheen, S. 1997. The beginnings of chronic airflow obstruction. *In* M.G. Marmot and M.E.J. Wadsworth, editors. Fetal and Early Childhood Environment. *Br. Med. Bull.* 53:58–70.
25. Barker, D. J. P. 1998. Mothers, Babies, and Disease in Later Life. British Medical Journal Publishing, London.
26. Wadsworth, M. E. J., S. L. Butterworth, R. J. Hardy, D. J. L. Kuh, M. Richards, C. Langenberg, W. S. Hilder, and M. Connor. 2003. Response and validity in a long-term prospective study. *Soc. Sci. Med.* 57:2193–2205.
27. Whitehouse, D. B., J. U. Lovegrove, and D. A. Hopkinson. 1989. Variation in alpha-1–antitrypsin phenotypes associated with penicillamine therapy. *Clin. Chim. Acta* 179:109–115.
28. Pauwels, R. A., A. S. Buist, P. M. Calverley, C. R. Jenkins, and S. S. Hurd. 2001. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for

Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am. J. Respir. Crit. Care Med.* 163:1256–1276.

29. Falaschetti, E., J. Laiho, P. Primatesta, and S. Purdon. 2004. Prediction equations for normal and low lung function from the Health Survey for England. *Eur. Respir. J.* 23:456–463.
30. Excoffier, L., and M. Slatkin. 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.* 12:921–927.
31. Von Ehrenstein, O. S., E. von Mutius, E. Maier, T. Hirsch, D. Carr, W. Schaal, A. A. Roscher, B. Olgemoller, T. Nicolai, and S. K. Weiland. 2002. Lung function of school children with low levels of alpha1-antitrypsin and tobacco smoke exposure. *Eur. Respir. J.* 19:1099–1106.

Appendix

TABLE A1. Value distributions of demographic variables used in the analysis (all available cases) [32]

	n, %	%
Social factors		
Sex		
Male	2,815	52.5
Female	2,547	47.5
Father's social class at 4 yr		
Nonmanual	1,834	34.2
Manual	3,528	65.8
Crowding at 2 yr (persons per room)		
Least crowded, 0.5	151	2.8
1	2,517	46.9
> 1.5	2,212	41.3
Unknown	482	9.0
Region of birth		
S. England	329	6
N. England	391	7
Yorkshire	380	7
N.W. England	562	10
N. Midlands	408	8
Midlands	473	9
E. England	373	7
London and S.E.	1,232	23
S.W. England	270	5
Scotland	656	12
Wales	288	5
Birth order		
First	2,272	42.4
Second	1,692	31.6
Third or later	1,390	25.9
Unknown	8	0.1
Parental educational attainment		
Primary	2,740	51.1
Secondary	985	18.4
Tertiary	522	9.7
Unknown	1,115	20.8
Developmental & nutritional factors		Mean
Birth weight (kg)	5,327	3.4
Unknown	35	–
Height at 2 yr (m)	4,010	0.85
Unknown	1,352	–
Weight at 2 yr (kg)	4,127	12.95
Unknown	1,235	–
		%
Infant feeding		
Not breastfed	1,133	21.1
Breast fed 1–4 mo	1,667	31.1
Breast fed > 4 mo	1,984	37.0
Unknown	578	10.8
Environmental factors		
Atmospheric pollution 0–2 yr		
Lowest quartile	870	16.2
Second quartile	1,210	22.6
Third quartile	1,081	20.2
Highest quartile	943	17.6
Unknown	1,258	23.5
Father's smoking		
No	608	11.3
Yes	2,355	43.9
Unknown	2,399	44.8
Mother's smoking		
No	1,609	30.0
Yes	1,354	25.3
Unknown	2,399	44.8

TABLE A2. Carrier status and infant lower respiratory infection by geographic region at birth

Region	Noncarriers *n*, %	S Carriers *n*, %	Z Carriers *n*, %	Infant Disease *n*, %
S. England	132, 82.5	18, 11.3	10, 6.3	44, 15.9
N. England	176, 84.2	22, 10.5	11, 5.3	108, 30.2
E. and W.Yorkshire	163, 84.0	27, 13.9	4, 2.1	89, 27.0
N.W. England	245, 88.8	23, 8.3	8, 2.9	163, 32.7
N. Midlands	198, 86.5	20, 8.7	11, 4.8	103, 29.3
Midlands	195, 90.3	19, 8.8	2, 0.9	121, 29.9
E. England	169, 82.0	22, 10.7	15, 7.3	60, 18.8
London and S.E.	554, 85.5	69, 10.6	25, 3.9	261, 24.6
S.W. England	123, 83.1	16, 10.8	9, 6.1	54, 22.0
Scotland	263, 89.8	19, 6.5	11, 3.8	100, 17.1
Wales	126, 83.4	17, 11.3	8, 5.3	72, 28.9
Total	2,344	272	114	1,175

Carrier status by region $\chi^2 = 29.2$; $P = 0.080$.
Region by infant lower respiratory infection $\chi^2 = 71.6$; $P = <0.001$.
Region by infant lower respiratory infection (S carriers only) $\chi^2 = 10.1$; $P = 0.431$.
Region by infant lower respiratory infection (Z carriers only) $\chi^2 = 8.3$; $P = 0.597$.