Original research

Chr15q25 genetic variant (rs16969968) independently confers risk of lung cancer, COPD and smoking intensity in a prospective study of high-risk smokers

Raewyn J Hopkins,[1] Fenghai Duan,[2] Greg D Gamble,[1] Caroline Chiles,[3] Alana Cavadino,[1] Paul Billings,[4] Denise Aberle,[5] Robert P Young[1]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/thoraxjnl-2020-214839).

[1]The University of Auckland Faculty of Medical and Health Sciences, Auckland, New Zealand
[2]Department of Biostatistics and Centre for Statistical Science, Brown University, Providence, Rhode Island, USA
[3]Department of Radiology, Wake Forest Baptist Medical Comprehensive Cancer Center, Winston-Salem, North Carolina, USA
[4]Natera Inc, Berkeley, California, USA
[5]Department of Radiological Sciences, University of California Los Angeles David Geffen School of Medicine, Los Angeles, California, USA

Correspondence to
Dr Robert P Young, The University of Auckland Faculty of Medical and Health Sciences, Auckland 1344, New Zealand; roberty@adhb.govt.nz

Received 17 March 2020
Revised 27 October 2020
Accepted 8 December 2020
Published Online First 8 January 2021

To cite: Hopkins RJ, Duan F, Gamble GD, *et al*. *Thorax* 2021;**76**:272–280.

ABSTRACT

Importance While cholinergic receptor nicotinic alpha 5 (CHRNA5) variants have been linked to lung cancer, chronic obstructive pulmonary disease (COPD) and smoking addiction in case–controls studies, their corelationship is not well understood and requires retesting in a cohort study.

Objective To re-examine the association between the CHRNA5 variant (rs16969968 AA genotype) and the development of lung cancer, relative to its association with COPD and smoking.

Methods In 9270 Non-Hispanic white subjects from the National Lung Screening Trial, a substudy of high-risk smokers were followed for an average of 6.4 years. We compared CHRNA5 genotype according to baseline smoking exposure, lung function and COPD status. We also compared the lung cancer incidence rate, and used multiple logistic regression and mediation analysis to examine the role of the AA genotype of the *CHRNA5* variant in smoking exposure, COPD and lung cancer.

Results As previously reported, we found the AA high-risk genotype was associated with lower lung function (p=0.005), greater smoking intensity (p<0.001), the presence of COPD (OR 1.28 (95% CI 1.10 to 1.49) p=0.0015) and the development of lung cancer (HR 1.41, (95% CI 1.03 to 1.93) p=0.03). In a mediation analyses, the AA genotype was independently associated with smoking intensity (OR 1.42 (95% CI 1.25 to 1.60, p<0.0001), COPD (OR 1.25, (95% CI 1.66 to 2.53), p=0.0015) and developing lung cancer (OR 1.37, (95% CI 1.03 to 1.82) p=0.03).

Conclusion In this large-prospective study, we found the CHRNA5 rs 16 969 968 AA genotype to be independently associated with smoking exposure, COPD and lung cancer (triple whammy effect).

Key messages

What is the key question?

► Is the Chromosome 15q25 genetic variant (rs16969968) association with lung cancer and smoking addiction independent of it's association with chronic obstructive pulmonary disease (COPD)?

What is the bottom line?

► Using a unique approach in a prospective study where genotyping and lung function were assessed at baseline, we have shown that a functional variant (rs16969968) encoding a subunit of the nicotinic receptor is independently associated with smoking intensity (cigarettes per day), spirometric defined COPD and development of lung cancer.

Why read on?

► This Chr 15q25 locus encodes a subunit of the nicotinic acetylcholine receptor that modifies the binding of nicotine in the airway and mediates the effects of inhaled nicotine in the lungs. Establishing this link in the largest prospective study to date directly implicates inhaled nicotine in the pathogenesis of both COPD and lung cancer which collectively account for about 50% of all deaths in current or former smokers. That nicotine, or its metabolites, are directly linked to the pathogenesis of lung cancer through binding its receptor challenges the prevailing assumption that inhaled nicotine is itself harmless.

INTRODUCTION

Large cross-sectional genome-wide association studies (GWAS) have consistently reported an association between the nicotinic acetylcholinereceptor alpha 5 subunit (CHRNA5) gene locus on chromosome 15q25 and lung cancer.[1–3] This locus has also been linked in cross-sectional studies to smoking exposure and addiction to smoking.[4–6] These findings have led to some debate as to whether the relationship between CHRNA5 and lung cancer is mediated through it effects with smoking, where increased risk of lung cancer is related to increased smoking exposure.[7 8] Soon after the original GWAS reported the lung cancer link, we reported that this locus was also associated with chronic obstructive pulmonary disease (COPD).[9] In our case–control study, we compared the frequency of the CHRNA5 variant rs16969968 in healthy smokers (normal lung function), with smokers with COPD and/or lung cancer relatively closely matched for age and smoking exposure history. This variant has been shown to change the structure and function of this subunit, consistent with it being a 'causative' disease

variant.[10 11] Our findings for COPD were quickly confirmed in a large COPD GWAS and subsequent case–control studies.[12] However, because COPD is also associated with lung cancer, the interactive effects of this locus on the development of lung cancer, independent of its effects on smoking or the development of COPD remain unclear.[7 13]

One way to explore the effect of this locus on the development of lung cancer is through a large prospective study, where current or former smokers at risk of lung cancer are followed prospectively after their smoking history and lung function were fully documented at baseline.[14 15] This reduces the bias that comes from using lung function performed exclusively for preoperative assessment; using lung cancers collected cross-sectionally and thus subject to survivor bias; or smoking history collected after the diagnosis of COPD or lung cancer was made.[16] A prospective study design including baseline spirometry also helps address the question of confounding, where biomarker associations made with lung cancer are potentially based on an unrecognised association with COPD.[17] The National Lung Screening Trial (NLST) is a large prospective study of 53 000 high-risk smokers followed over a mean of 6.4 years.[18] Among this cohort, there were 9270 Non-Hispanic whites, including 3500 current or former smokers with airflow limitation (35%) at baseline, and 380 lung cancers diagnosed during the follow-up period. Preliminary results of this study have been reported in abstract form.[19]

METHODS

Subjects

This is a secondary data analysis of the NLST. The recruitment and study design of this trial, involving 53 452 screening participants yielding 2058 histology confirmed lung cancers, has been described elsewhere.[18] In the American College of Radiology Imaging Network (ACRIN) subcohort of the NLST, participants from 23 centres agreed to undergo baseline prebronchodilator spirometry (NLST-ACRIN Biorepository Cohort, N=18 842) and, for a subgroup (figure 1), blood sampling for biomarker analysis (N=10 054). Demographic data, including history of premorbid disease, were collected through an extensive questionnaire and shows this NLST-ACRIN cohort to be highly representative of the full NLST cohort.[14 15] From the total group of 10 054 (including all ethnicities), we analysed genomic data for Non-Hispanic whites comprising 9270 high-risk smokers from which 380 lung cancers were diagnosed during the study follow-up (figure 1).

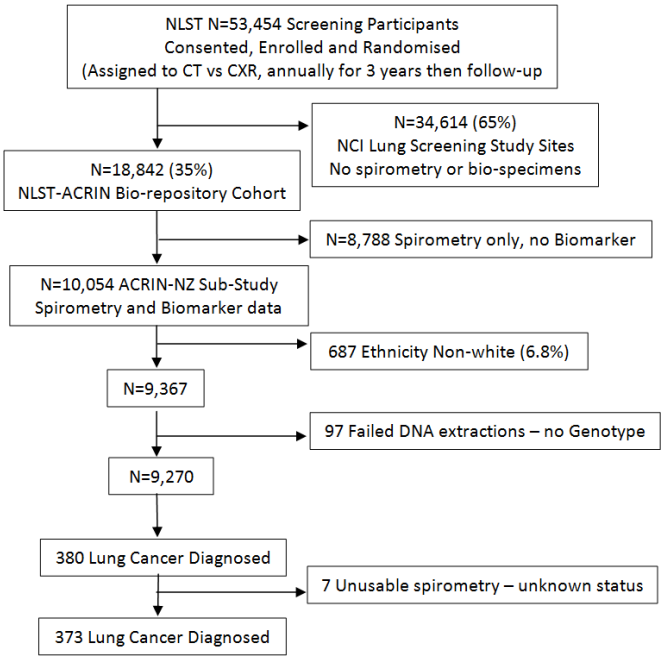

Figure 1 CONSORT figure of the genetic study subgroup from the NLST. ACRIN, American College of radiology, imaging network; CONSORT, Consolidated Standards of Reporting Trials; NCI, National cancer Institute; NLST, National Lung Screening Trial; NZ, New Zealand.

Pulmonary function testing

In the NLST-ACRIN cohort, prebronchodilator spirometry was measured at baseline screening (T0) in the majority of participants meeting previously published criteria.[18] The spirometry was measured by trained staff using a Spiropro spirometer (eResearchTechnology, Germany). The severity of airflow limitation was defined according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria grades 1–4 (www.GOLD.org accessed 2 February 2020). Those with no airflow limitation were further subgrouped into those with normal lungs ('Resistant Smokers') and Restrictive Spirometry, where the latter is defined as forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) ≥0.70 and FEV_1%predicted <80% (http://www.copdgene.org/study-design accessed 2 February 2020). For comparative purposes[9 13] in this study, COPD was defined as GOLD 2–4 grade.

Lung cancer outcomes

Lung cancer cases included all those diagnosed during the trial (N=380), whether screen or non-screen detected (interval), or prevalent (diagnosed at T0 or during the first year) or incident lung cancers (diagnosed during subsequent years T1–T6) or at post-mortem.[18] All lung cancer cases were confirmed on histological sampling according to accepted international classification criteria. Lung function results and mortality outcomes were available for 373 of the 380 lung cancer cases (98% of total). The NLST was terminated early when the endpoint of a 20% reduction in lung cancer-specific mortality in the CT arm, relative to the chest X-ray (CXR) arm, was reached with a mean follow-up of 6.4 years. Cause of death was a primary outcome for the NLST, and was ascertained through review of clinical records and death certification.

Genotyping

Genomic DNA was extracted from buffy coat samples using standard salt-based methods and purified genomic DNA was aliquoted (10 ng·µ/L concentration) into 384-well plates. Samples were genotyped for rs16969968 of the CHRNA 5 gene using the Sequenom system (Sequenom Autoflex Mass Spectrometer and Samsung 24 pin nanodispenser) by Agena (San Diego, USA). The Sequenom sequences were designed in house by Agena with amplification and separation methods (iPLEX, www.sequenom.com) as previously described.[20]

Statistical analysis

Differences according to CHRNA5 genotype were examined using both allelic (0, 1 or 2 'A' alleles), genotype (AA vs AG vs GG) and recessive (AA vs AG/GG combined) models. χ^2 tests, test for trend or Fisher's exact test (for small-cell counts) were used to compare categorical variables, and t-tests or Kruskal-Wallis tests (for non-normally distributed variables) were used to compare continuous variables by genotype. Lung function measures at baseline were assessed using robust linear regression to account for outlying values and adjust for pack-years, age and sex. Prevalence rates and 95% CIs for lung cancer were

calculated per 1000 person-years. Cox proportional hazards models were used to compare survival adjusted for pack-years, with estimates presented as HRs and 95% CIs. The proportional hazards assumption was verified. Multiple logistic regression modelling was used to model odds with and without adjustment as indicated. For models of association with COPD status and lung cancer diagnosis, allelic (log additive) and recessive models were compared with the general genotype model (AA vs AG vs GG) using a likelihood ratio test (LRT), Akaike information criterion (AIC) and Bayesian information criterion (BIC) metrics, to assess which model provided the best fit. Mediation analysis was performed using the CAUSALMED procedure of SAS (SAS V.9.4, SAS Institute) fitting a binomial distribution with logit link function to the binary outcome lung cancer outcome as function of the direct and indirect effect of the binary CHRNA5 recessive variable with dichotomous GOLD group as a mediator and pack-years >50 years as a categorical covariate. Total effect decompositions were estimated. Statistical significance was defined as a two tailed $p<0.05$. All analyses were performed using SAS (V.9.4) or STATA statistical software (V.16StataCorp).

RESULTS

Of the Non-Hispanic whites (N=9367, 93%) of the total NLST-ACRIN cohort, genotype data for the CHRNA5 nicotinic receptor polymorphism (rs16969968) was available for 9270 subjects (99%). In table 1, the demographic variables and premorbid self-reported diseases were compared according to the AA (homozygous minor allele, 12%), AG (heterozygous, 46%) and GG (homozygous major allele, 42%). These genotype

Table 1 Non-Hispanic whites:CHRNA rs16969968 genotypes—baseline demographic and comorbidity data according to allelic and recessive regression models

CHRNA rs16969968 N=9270*	AA	AG	GG	Allelic P value	Recessive P value
Genotype (% total)	1153 (12.4%)	4258 (45.9%)	3859 (41.6%)	–	–
Demographics					
Age (years)	61.6±5.0	61.8±5.1	61.8±5.1	0.18	0.17
Male	655 (56.8%)	2416 (56.7%)	2198 (57.0%)	0.98	0.98
Family history of lung cancer	276 (23.9%)	1049 (24.6%)	881 (22.8%)	0.16	0.91
Self-report COPD (composite)	255 (22.1%)	897 (21.1%)	752 (19.5%)	0.08	0.16
Smoking history					
Current smoker	561 (48.7%)	2061 (48.4%)	1839 (47.6%)	0.74	0.70
Pack-years	59.7±25.4	56.6±23.3	54.7±22.5	**<0.001**	**<0.001**
Cigarettes per day	30.0±11.8	28.4±11.1	27.5±10.8	**<0.001**	**<0.001**
Years quit	3.6±5.0	3.8±5.1	3.7±5.1	0.86	0.60
Smoking duration (years)	40.2±7.5	40.4±7.4	40.3±7.5	0.69	0.58
Body composition					
Body mass index (kg/m^2)	27.7±5.0	27.8±5.1	28.0±5.1	0.52	0.31
Weight (kg)	82.0±18.0	82.4±17.7	83.0±18.1	0.45	0.23
Height (cm)	171.6±9.9	171.8±9.9	172.0±9.9	0.89	0.34
Education level					
High school or less	329 (28.5%)	1185 (27.8%)	1127 (29.2%)	0.27	0.16
Post high school/some college	389 (33.7%)	1549 (36.4%)	1361 (35.3%)		
College grad/postgrad/professional	412 (35.7%)	1407 (33.0%)	1274 (33.0%)		
Other/unknown	23 (2.0%)	117 (2.7%)	97 (2.5%)		
Pre morbid disease (self-report)					
COPD	106 (9.2%)	338 (7.9%)	230 (6.0%)	**<0.001**	**0.007**
Chronic bronchitis	123 (10.7%)	469 (11.0%)	437 (11.3%)	0.80	0.62
Emphysema	116 (10.1%)	409 (9.6%)	328 (8.5%)	0.13	0.28
Adult asthma	85 (7.4%)	298 (7.0%)	253 (6.6%)	0.56	0.46
Pneumonia	336 (29.1%)	1190 (27.9%)	1074 (27.8%)	0.67	0.38
Heart disease	152 (13.2%)	578 (13.6%)	511 (13.2%)	0.89	0.83
Hypertension	407 (35.3%)	1534 (36.0%)	1365 (35.4%)	0.80	0.78
Stroke	31 (2.7%)	123 (2.9%)	119 (3.1%)	0.75	0.58
Diabetes	111 (9.6%)	379 (8.9%)	354 (9.2%)	0.74	0.51
Any cancer history	54 (4.7%)	181 (4.2%)	151 (3.9%)	0.48	0.35

*Minus subjects who failed genotyping n=97 (1%). Allelic AA versus AG versus GG. Recessive AA versus AG/GG. Composite COPD= 'yes' to doctor diagnosed 'COPD' or 'emphysema' or 'chronic bronchitis' or 'adult asthma'.
COPD, chronic obstructive pulmonary disease.

Table 2 Non-Hispanic whites: CHRNA rs16969968 relationship with lung function and COPD status according to allelic, recessive and genotype models using logistic regression. (A) baseline lung function, gold grade AL and airway phenotypes; (B) Unadjusted and adjusted effect estimates for association with baseline lung function

(A) ACRIN Non-Hispanic whites					
CHRNA rs16969968	**AA**	**AG**	**GG**	**Allelic Model**	**Recessive Model**
Genotype N=9270 (% total cohort)	N=1153 (12.4%)	N=4258 (45.9%)	N=3859 (41.6%)	P value	P value
Lung Ffunction (baseline)					
FEV1/FVC (%, mean ±SD)	70.1 ± 11.0	70.6 ± 10.9	71.7 ± 10.4	<0.001*	0.014*
FEV1 % predicted (mean± SD)	79.7 ± 20.9	80.7 ± 20.0	82.5 ± 19.8	<0.001*	0.005*
FVC % predicted (mean ± SD)	86.0 ± 18.4	86.6 ± 17.3	87.2 ± 17.1	0.023*	0.072*
AL (GOLD grade)†					
GOLD 1, N=814	94 (8.15%)	405 (9.51%)	315 (8.16%)	0.062	0.95
GOLD 1–4, N=3186	433 (37.55%)	1541 (36.19%)	1,212 (31.41%)	<0.001	0.008
GOLD 2–4 (FEV1 ≤80%), N=2372	339 (29.40%)	1136 (26.68%)	897 (23.24%)	<0.001	<0.001
GOLD 3–4 (FEV1 ≤50%), N=608	91 (7.89%)	296 (6.95%)	221 (5.73%)	<0.001	0.017
COPD GOLD 2–4 – N (% by genotype)	339 (29%)	1136 (27%)	897 (23%)	-	-
OR unadjusted (95% CI)	1.41 (1.21 to 1.64)	1.23 (1.11 to 1.36)	Ref	1.20 (1.12 to 1.29)	1.26 (1.10 to 1.45)
P value	<0.001	<0.001		<0.001	<0.001
OR adjusted‡ (95% CI)	1.32 (1.13 to 1.54)	1.20 (1.08 to 1.33)	Ref	1.16 (1.08 to 1.25)	1.19 (1.04 to 1.38)
P value	<0.001	0.001		<0.001	0.014
No AL (AL)					
No AL total, N=5927 (%)	692 (60.02%)	2655 (62.35%)	2580 (66.86%)	-	-
Restrictive Spirometry, N=1460	172 (11.78%)	675 (46.23%)	613 (41.99%)	0.280†	0.885†
Resistant smokers, N=4467	520 (11.64%)	1980 (44.33%)	1967 (44.03%)		

(B) Model	Adjustment	FEV1/FVC, % Estimate (95% CI)	P value	FEV1 % predicted Estimate (95% CI)	P value	FVC, % predicted Estimate (95% CI)	P value
Allelic	None	−0.80 (−1.08 to −0.51)	<0.001	−1.64 (−2.24 to −1.05)	<0.001	−0.73 (−1.21 to −0.25)	0.003
	Pack-years	−0.60 (−0.88 to −0.31)	<0.001	– 1.24 (−1.83 to −0.65)	<0.001	−0.49 (−0.96 to −0.02)	0.042
	Pack-years, age, sex	−0.67 (−0.95 to −0.39)	<0.001	−1.31 (−1.89 to −0.73)	<0.001	−0.55 (−1.02 to −0.08)	0.023
Recessive	None	−0.91 (−1.50 to −0.32)	0.003	−2.20 (−3.42 to −0.98)	<0.001	−1.17 (−2.12 to −0.19)	0.019
	Pack-years	−0.60 (−1.18 to −0.02)	0.044	−1.55 (−2.75 to −0.35)	0.011	−0.76 (−1.73 to 0.21)	0.124
	Pack-years, age, sex	−0.72 (−1.29 to −0.15)	0.014	−1.71 (−2.91 to −0.52)	0.005	−0.89 (−1.85 to 0.08)	0.072

No spirometry for 152 subjects.
*P value adjusted for pack-years, age and gender.
†Comparison is for restrictive spirometry versus resistant smokers within the no AL group.
‡OR adjusted for pack-years (see figure 2).
ACRIN, American College of Radiology Imaging Network; AL, airflow limitation; COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1s; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

frequencies accord with Hardy-Weinberg equilibrium and were consistent with the published frequencies in other Caucasian populations.[1–6 9 12] Using this method, we have previously shown 100% concordance in genotyping with an SNP in complete linkage disequilibrium.[19 21] We found that across the three genotypes, age, gender, family history for lung cancer, smoking status, and years quit, smoking duration, body mass index and educational level were not significantly different (table 1). We found the A allele (allelic model) and AA genotype (recessive model) were associated with higher pack-years (p<0.001) and higher cigarettes per day (p<0.001) with no differences in years quit or years smoked. A similar finding was found for self-reported COPD but not for other premorbid diseases at baseline (table 1).

Comparable to our first study,[9] on comparing lung function and COPD severity (GOLD grade) in table 2A, we found that the AA genotype (recessive model) was associated with lower FEV_1%predicted (p=0.005), greater COPD (GOLD 2–4) frequency (p<0.001) but not GOLD 1 COPD. Indeed, the AA frequency increased as the severity of airflow limitation increased from 'Resistant smokers' (11.6%), GOLD 1 (11.5%), GOLD 2–4 (14.3%) and GOLD 3–4 (15.0%) (table 2A). Moreover, the fully adjusted effect estimates were greatest in the recessive model and most specific for airflow limitation (least significant for FVC% predicted) (table 2B). However, on comparing the odds of COPD (GOLD 2–4) adjusted for smoking intensity, there was evidence the general genotype model fit the data better than a recessive model (LRT p=0.0005), and further that the allelic model provided the best fit (lowest AIC and BIC values), with no evidence that the genotype model was an improvement (LRT vs allelic model, p=0.387) (online supplemental table 1A). Adjusted for smoking intensity, there was a 1.16 (95% CI 1.08 to 1.25, p<0.001) increase in the odds of COPD for each additional A allele (table 2A, figure 2). Collectively these findings suggest a log additive (allelic) model is most consistent for COPD (GOLD 2–4) at this locus (table 2 and figure 2). In those with no airflow limitation, consisting of those with 'restrictive'

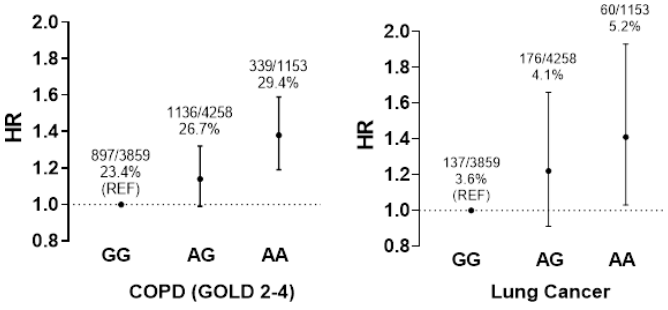

Figure 2 OR for COPD (GOLD 2–4) and HR for lung cancer, according to rs16969968 CHRNA genotype after adjustment for smoking pack-years (±95% CI) and referenced against the GG genotype. COPD, chronic obstructive pulmonary disease; CHRNA, cholinergic receptor nicotinic; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

or 'resistant' spirometric subgroups, there were no allele or genotype differences (table 2A).

Consistent with our previous study,[9] and after adjustment for smoking, the AA genotype was also associated with a greater prevalence and incidence of lung cancer overall (table 3, figure 2). In the current study, where lung cancer was identified prospectively, the AA genotype relative to the GG genotype had an adjusted HR=1.39 (95% CI 1.02 to 1.88, p=0.034) and for the AG genotype an adjusted HR=1.14 (0.91 to 1.43, p=0.243) (table 3, figure 2). Adjusted for smoking intensity, there was a 1.17 (95% CI 1.01 to 1.36, p=0.035) increase in the hazard of a lung cancer (overall) for each additional A allele (table 3, online supplemental table 1B). On comparing the risk of a lung cancer using the LRT, and adjusted for smoking intensity, we found no evidence that the general genotype model fitted the data better than a recessive model (LRT p=0.242) or an allelic model (LRT p=0.770). While AIC and BIC values were slightly lower in the allelic model (online supplemental table 1B), we note that after stratifying lung cases for COPD (table 4), the lung cancer with COPD data best fit an allelic model while for lung cancer alone the data best fit a recessive model (based on LRT analysis and AIC results, data not shown). This is consistent with our finding that the AA genotype, but not the AG genotype, was significantly associated with lung cancer relative to GG (figure 2). We found no major differences in lung cancer characteristics or demographic variables across the three genotypes (table 3). While the allelic (log additive) model best fit the data for COPD (tables 2 and 4, figure 2 and online supplemental table 1A), we suggest the recessive model appears to best fit the data for lung cancer, especially when those with COPD are excluded (tables 3 and 4, online supplemental table 1B). In table 4, relative to the referent group (AG=45%), the excess in AG genotype is a feature of COPD GOLD 2–4 (AG=48%) and Lung cancer with COPD (50%), but not for those with lung cancer alone (44%).

In a multiple logistic regression, where the AA genotype was compared with the other genotypes, age and smoking duration (grouped in 5-year bands), cigarettes per day (grouped as five per day bands) and FEV_1/FVC (grouped in 5% bands), independently contributed to lung cancer risk (figure 3). In a mediation analysis, we confirmed that smoking (50+ pack-years), COPD (GOLD 2–4) and the CHRNA5 AA genotype all contributed independently to the risk of lung cancer (figure 4). CHRNA5 AA genotype contributed directly (independently) 9.0% of the effect (p=0.03).

DISCUSSION

This study confirms our earlier findings that the CHRNA5 polymorphism (rs16969968) is independently associated with smoking, the presence of airflow limitation (COPD) and lung cancer.[9] We found a dose–response relationship between the CHRNA5 A allele with pack-years, cigarettes per day, lung function (%predFEV_1) and airflow limitation (FEV_1/FVC), with a log-additive allelic model providing the best fit for COPD. We also found evidence for a linear relationship between the A allele and self-reported COPD, the presence of COPD defined by spirometry and the incidence of lung cancer. While the AA genotype (recessive model) was also associated with smoking history, lung function, airflow limitation (COPD) and lung cancer incidence, the allelic and recessive models were a comparable fit for lung cancer. These findings suggest the CHRNA5 locus is not only associated with smoking exposure but links COPD and lung cancer at a molecular genetic level. Such a hypothesis has been debated for over 40 years based on early genetic epidemiological studies and raises a number of important issues.[22 23]

First the genetic overlap between COPD and lung cancer supports the hypothesis that these diseases share pathogenic pathways.[24] This finding has been made possible by measuring lung function routinely across high-risk smokers followed prospectively in a large cohort study.[18] Such an approach is unique in lung cancer genetics where most studies are cross-sectional case–control in design and make no provision for the mediating effects of coexisting airflow limitation.[1–3 25] While large studies confirm the association between this locus and lung cancer (or smoking exposure), the mediating role of COPD has not been assessed in the same study population.[25] The rs16969968 polymorphism has been shown to have a functional effect on the activity of the nicotinic acetylcholine receptor found in both the respiratory epithelium where it modulates inflammatory pathways and in the brain where it modulates addiction pathways.[10 11 26] Nicotine replaces acetylcholine in receptor activation indicating that nicotine itself has a pathogenic role in COPD and lung cancer, not just isolated to promoting addiction to cigarette smoking. This is relevant because there is a growing body of literature suggesting exposure to inhaled nicotine, such as from vaping, is directly pathogenic in the lungs.[27 28] It also has relevance in the regulation of nicotine levels in its various inhaled forms. The findings of this study fuel further concerns about the long-term dangers of chronic nicotine inhalation.

Second, this study highlights the close relationship between COPD and lung cancer from a genetic epidemiological perspective. Given the well-publicised genetic studies of lung cancer make little provision for co-existing COPD,[1–3] there is the possibility that some of the reported associations (and subsequent genetic modelling) result from a confounding or mediating effect from the disproportional presence of COPD in the cases (vs controls) in these retrospective studies. Furthermore we suggest that genetic models may be different in lung cancer according to the presence or absence of COPD. We have demonstrated this previously with regards to the glutathione S-transferase M polymorphism linked to lung cancer in a large meta-analysis but shown to be confounded by unrecognised COPD.[17] This close relationship has relevance to the use of genetic data in risk-based approaches to targeted interventions such as CT screening for lung cancer.[29 30] We have shown that those with normal lung function or only mild COPD gain the most from lung cancer screening.[31] In contrast, those with severe or very severe COPD achieve little or no reduction in lung cancer death from screening.[31] This is because COPD is associated with more

Table 3 Non-Hispanic Whites: CHRNA rs16969968 genotypes and lung cancer (LC) outcomes according to allelic, recessive and genotype models using COX proportional hazards models

ACRIN non-Hispanic whites					
CHRNA rs16969968	**AA**	**AG**	**GG**	**Allelic Model P value**	**Recessive Model P value**
LC diagnosed n=373	60	176	137		
Prevalence per/100 screened	5.2%	4.1%	3.6%	0.013	0.029
Incidence/1000 person-years (95% CI)	8.5 (6.6 to 10.9)	6.7 (5.8 to 7.8)	5.8 (4.9 to 6.8)	–	–
HR unadjusted (95% CI)	1.48 (1.09 to 2.01)	1.17 (0.93 to 1.46)	ref	1.21 (1.04 to 1.40)	1.36 (1.03 to 1.80)
P value	0.011	0.177		0.012	0.028
HR adjusted (95% CI)†	1.39 (1.02 to 1.88)	1.14 (0.91 to 1.43)	ref	1.17 (1.01 to 1.36)	1.29 (0.98 to 1.70)
P value	0.034	0.243		0.035	0.070
Time to LC diagnosis or end of follow-up—mean years (SD)	6.02 (1.33)	6.08 (1.25)	6.09 (1.24)	–	–
LC according to COPD status					
With COPD (GOLD 2–4)	23 (15.2%)	75 (49.7%)	53 (35.1%)	0.630	0.793
Without airflow limitation	29 (16.3%)	78 (43.8%)	71 (39.9%)		
LC histology (N=% grp LCdx)					
1 small cell	7 (11.7%)	25 (14.2%)	17 (12.4%)	0.60	0.26
2 squamous cell	16 (26.7%)	38 (21.6%)	34 (24.8%)		
3 adenocarcinoma	18 (30.0%)	64 (36.4%)	51 (37.2%)		
4 BAC	8 (13.3%)	11 (6.25%)	15 (10.9%)		
5 large cell	3 (5.0%)	2 (1.1%)	2 (1.5%)		
6 non-small cell	8 (13.3%)	36 (20.5%)	17 (12.4%)		
7 other	–	–	1 (0.73%)		
LC stage at Dx (N=% grp LCdx)					
1 stage I	35 (58.3%)	72 (40.9%)	57 (41.6%)	0.27	**0.04**
2 stage II	–	10 (5.7%)	8 (5.8%)		
3 stage III	10 (16.7%)	36 (20.5%)	31 (22.6%)		
4 stage IV	14 (23.3%)	56 (31.8%)	37 (22.6%)		
5 occult carcinoma/unk	1 (1.7%)	2 (1.1%)	4 (2.9%)		
Surgery					
Surgery LC-yes (N=% total LCdx)	36 (60%)	84 (47.7%)	77 (56.2%)	0.96	0.22
Screening interval (N=% grp LCdx)					
T0-T2	40 (66.7%)	105 (59.7%)	87 (63.5%)	0.89	0.47
T3-T6	20 (33.3%)	71 (40.3%)	50 (36.5%)		
LC detection (N=% grpLCdx)					
Screen-detected	32 (53.3%)	85 (48.3%)	75 (54.7%)	0.79	0.49
Missed	5 (8.3%)	16 (9.1%)	10 (7.3%)		
Interval	3 (5%)	4 (2.3%)	2 (1.5%)		
Follow-up	20 (33.3%)	71 (40.3%)	50 (36.5%)		
Other demographics of LCdx					
% Male	25 (41.6%)	104 (59.1%)	77 (56.2%)	0.16	**0.02**
% Current smoker	34 (56.7%)	106 (60.2%)	72 (52.6%)	0.39	0.98
Family history of LC –(yes)	13 (21.7%)	49 (27.8%)	35 (25.5%)	0.74	0.40
Self-reported COPD (composite)	21 (35%)	64 (36%)	31 (23%)	**0.03**	0.54
Randomised CT	30 (50%)	97 (55.1%)	72 (52.6%)	0.89	0.58

*Model adjusted for pack-years.
†See figure 2.
ACRIN, American College of Radiology Imaging Network; COPD, chronic obstructive pulmonary disease.

aggressive lung cancer, more late stage disease, lower surgical rates and higher rates of deaths from cardiovascular and respiratory causes.[31 32]

Third, understanding the genetic basis of lung cancer requires some consideration of the contributory effects of the genetic basis of COPD. Indeed the heritability of COPD, estimated to be 50%–70%, is two to threefold greater than for lung cancer, estimated to be about 20%–30%. Between 50% and 80% of smokers with lung cancer have COPD depending on how the latter is defined.[9 13 24] This means that the genetic basis of lung cancer may include a large contribution from a genetic overlap with the tendency for COPD. The only way to successfully tease

Table 4 Non-Hispanic whites only: CHRNA rs16969968 genotype frequencies according to AL, lung cancer (LC), LC subphenotypes according to the allelic and recessive models using logistic regression

ACRIN Non-Hispanic whites only (N=9270)					
CHRNA rs16969968	**AA**	**AG**	**GG**	**Allelic P value***	**Recessive P value†**
Genotype N=9270 (% total)	N=1153 (12.4)	N=4258 (45.9)	N=3859 (41.6%	–	–
No AL					
No AL, N=5921 (% total)	692 (11.7)	2653 (44.8)	2576 (43.5)	–	–
No AL and no LC, N=5749 (% total)‡ (referent)	663 (11.5)	2577 (44.8)	2509 (43.6)	–	–
% COPD (FEV_1/FVC <0.70)					
GOLD 1, N=814 (% total)	94 (11.5)	405 (49.8)	315 (38.7)	0.101	0.763
GOLD 2–4 (FEV_1≤80%), N=2372 (% total)	339 (14.3)	1136 (47.9)	897 (37.8)	<0.001	0.012
LC					
Total LC diagnosis, n=373 (% total)	60 (16.1)	176 (47.2)	137 (36.7)	0.008	0.033
Squamous cell LC, N=88 (% total)	16 (18.2)	38 (43.2)	34 (38.6)	0.207	0.126
Adenocarcinoma LC, N=133 (% total)	18 (13.5)	64 (48.1)	51 (38.3)	0.356	0.688
LC and COPD status					
LC and GOLD 2–4, N=151 (% total)	23 (15.2)	75 (49.7)	53 (35.1)	0.077	0.354
LC and no AL, N=178 (% total)	29 (16.3)	78 (43.8)	71 (39.9)	0.163	0.092

*AA versus AG versus GG.
†AA versus AG/GG.
‡No AL 'healthy smokers' with no LC diagnosis during follow-up is the reference group for all subsequent comparisons. All analyses are adjusted for pack-years.
ACRIN, American College of Radiology Imaging Network; AL, airflow limitation; FEV1, forced expiratory volume in 1s; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

out the genetic basis of lung cancer is to account for the presence of COPD and its contribution to risk (and genetic modelling), as we have done in this study. This is only possible where lung function data are available in the same cohort of people getting lung cancer. The current study is the only cohort study of this type that we are aware of. In a case–control study, VanderWeele *et al* used mediation analysis to confirm this CHRNA5 locus was linked to lung cancer independent of smoking exposure.[33]

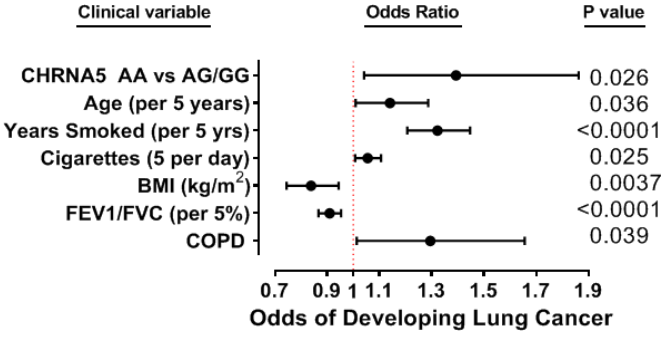

Figure 3 Multiple logistic regression analysis for risk of lung cancer according to genotype, smoking, BMI, lung function and presence of COPD (GOLD 2–4). BMI, body mass index; COPD, chronic obstructive pulmonary disease; CHRNA5, cholinergic receptor nicotinic alpha 5; FEV1, forced expiratory volume in 1s; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

Not only do our mediation results concur with the VanderWeele study, we have expanded their findings by showing the independent pathways linking this locus with lung cancer may in part be related to pathogenetic pathways underlying COPD.[34 35] This is important, as suggested above, when it comes to using genetic data in assessing risk of lung cancer for targeted interventions

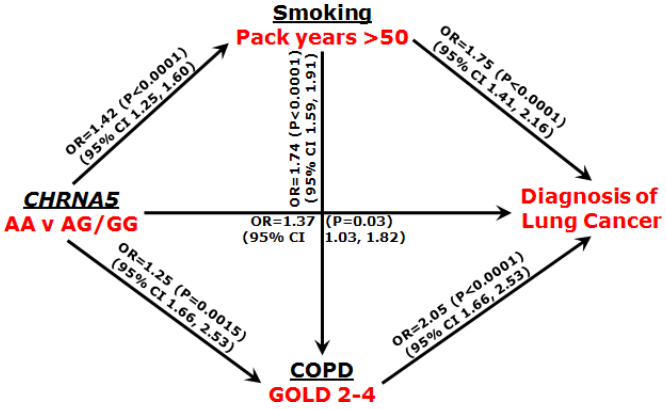

Figure 4 Mediation analysis comparing the relative contributions of rs16969968 AA genotype, pack-years and presence of COPD (GOLD 2–4) on the risk of developing lung cancer. CHRNA5, cholinergic receptor nicotinic alpha 5; COPD, chronic obstructive pulmonary disease; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

such as screening.[29 31] We have shown here that the CHRNA5 polymorphism (rs16969968) mediated risk of both COPD and lung cancer independently, although by different genetic models (allelic and recessive respectively). We have also previously found that some genetic variants confer risk of COPD only, lung cancer in the absence of COPD[13 20] and greater lung cancer lethality.[36] This is of particular relevance because smokers with normal lung function, or only mild COPD, achieve significantly better reductions in lung cancer deaths with CT-based screening relative to CXR screening in the NLST.[31]

The current study has several strengths and limitations. First, the data linking the CHRNA5 polymorphism (rs16969968) polymorphism to airflow limitation is based on cross-sectional data from a cohort of heavy smokers rather than prospective data from a population more representative of the wider non-smoking and smoking community. For example, HUNT study.[7] That said, while the study population we used was a smoking cohort of high risk heavy smokers, it is representative of screening populations and provides a more stringent test of the smoking-by-gene interaction. This contrasts with previous case-control studies that included lighter smokers of lower overall risk to identify comparable associations.[1–6 25] Second, we have only investigated one polymorphic locus and not considered the relevance of other variants in this chromosomal region (eg, iron-responsive element-binding protein 2, IREB2).[37] That the rs 16969968 polymorphism has been shown to have functional effects on gene expression and protein function,[10 11 34 35] and a consistent relationship between the presence of the A allele and increased risk, supports the possibility that this is indeed an important disease-causing polymorphism. Further studies will be required to examine possible epistatic effects with neighbouring or distant variants including from other candidate genes (eg, IREB2, PSMA4). While residual confounding and incomplete modelling might remain an issue in our study, we have at least attempted to account for the complex interplay between smoking exposure, COPD and lung cancer through our mediation analysis. Third due to relatively small sample sizes, type 1 error (false positive result) remains a possible explanation for our findings. However, the consistency of the association between this variant and lung cancer, and consistency in the magnitude of the effect size with other larger studies is reassuring. We have confirmed the association using both a genotype (table 3) and phenotype (table 4) approach. Fourth, McKay *et al* recently reported that in a large case–control study, this variant was associated specifically to the squamous cell histological subgroup of lung cancer.[25] While our study is underpowered to explore subgroup associations, our analysis supports this finding (table 4). However, this finding by McKay might be explained through mediating or confounding effects of COPD. A major strength of this study is the validation of a molecular marker for lung cancer risk in a population for whom assessing lung cancer risk has clinical relevance, in this case for lung cancer screening and its outcomes.[38]

In summary, we have used a prospective cohort study to confirm that the A allele and/or AA genotype of the CHRNA5 (rs16969968) polymorphism is independently associated with smoking exposure, presence of significant airflow limitation and risk of developing lung cancer. The study design we have used is unique in confirming this complex inter-relationship and highlights the need to consider the mediating effect of COPD in lung cancer genetics. Our findings support those of others suggesting the nicotine receptor, and by inference inhaled nicotine exposure, may have direct effects on the development of COPD and lung cancer. Our findings confirm our previous hypothesis that COPD and lung cancer are linked at a molecular genetic level.

Contributors Author contribution: RJH, FD, CC, PB, DA and RPY contributed to the conception and design; acquisition, analysis and interpretation; drafting and review for important intellectual content and final approval of the manuscript. GDG and AC contributed to biostatistical analysis and final approval.

Funding This study was funded by a grant from Johnson and Johnson and grants U01-CA-80098 and U01-CA-79778 to the American College of Radiology Imaging Network (ACRIN).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information. All clinical data is deidentified. The genetic data are reported in the manuscript.

REFERENCES

1 Hung RJ, McKay JD, Gaborieau V, *et al*. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on *15q25*. *Nature* 2008;452:633–7.
2 Thorgeirsson TE, Geller F, Sulem P, *et al*. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 2008;452:638–42.
3 Amos CI, Wu X, Broderick P, *et al*. Genome-Wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;40:616–22.
4 Saccone NL, Wang JC, Breslau N, *et al*. The CHRNA5-CHRNA3-CHRNB4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. *Cancer Res* 2009;69:6848–56.
5 Saccone NL, Culverhouse RC, Schwantes-An T-H, *et al*. Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. *PLoS Genet* 2010;6:e1001053.
6 Chen L-S, Baker T, Hung RJ, *et al*. Genetic risk can be decreased: Quitting smoking decreases and delays lung cancer for smokers with high and low *CHRNA*5 risk genotypes - A meta-analysis. *EBioMedicine* 2016;11:e00734:219–26.
7 Gabrielsen ME, Romundstad P, Langhammer A, *et al*. Association between a 15q25 gene variant, nicotine-related habits, lung cancer and COPD among 56,307 individuals from the HUNT study in Norway. *Eur J Hum Genet* 2013;21:1293–9.
8 Wang Y, Broderick P, Matakidou A, *et al*. Chromosome 15q25 (*CHRNA3-CHRNA5*) variation impacts indirectly on lung cancer risk. *PLoS One* 2011;6:e19085.
9 Young RP, Hopkins RJ, Hay BA, *et al*. Lung cancer gene associated with COPD: triple whammy or possible confounding effect? *Eur Respir J* 2008;32:1158–64.
10 Wang JC, Cruchaga C, Saccone NL, *et al*. Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. *Hum Mol Genet* 2009;18:3125–35.
11 Smith RM, Alachkar H, Papp AC, *et al*. Nicotinic α5 receptor subunit mRNA expression is associated with distant 5′ upstream polymorphisms. *Eur J Hum Genet* 2011;19:76–83.
12 Pillai SG, Ge D, Zhu G, *et al*. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* 2009;5:e1000421.
13 Young R, Hopkins R, Etzel C, *et al*. Genetics of lung cancer susceptibility and COPD. *Lancet Oncol* 2011;12:622–3.
14 Young RP, Duan F, Chiles C, *et al*. Airflow limitation and histology shift in the National lung screening trial. The NLST-ACRIN cohort substudy. *Am J Respir Crit Care Med* 2015;192:1060–7.
15 Hopkins RJ, Duan F, Chiles C, *et al*. Reduced expiratory flow rate among heavy smokers increases lung cancer risk. results from the National lung screening Trial-American College of radiology imaging network cohort. *Ann Am Thorac Soc* 2017;14:392–402.
16 Mohamed Hoesein FA, de Jong PA. Personalized lung cancer screening: the value of spirometry and emphysema as risk modifiers. *Ann Transl Med* 2016;4:293.
17 Young RP, Hopkins RJ, Hay BA, *et al*. *GSTM1* null genotype in COPD and lung cancer: evidence of a modifier or confounding effect? *Appl Clin Genet* 2011;4:137–44.
18 National Lung Screening Trial Research Team, Aberle DR, Adams AM, *et al*. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395–409.
19 Hopkins RJ, Young RP, Duan F, *et al*. Chromosome 15q25 locus and genetic susceptibility to lung cancer, COPD and smoking:Triple whammy effect in the ACRIN-NLST cohort sub-study (n=10,054). *Am J Respir Crit Care Med* 2017;195:A7052.
20 Young RP, Hopkins RJ, Whittington CF, *et al*. Individual and cumulative effects of GWAS susceptibility loci in lung cancer: associations after sub-phenotyping for COPD. *PLoS One* 2011;6:e16476.
21 Young RP, Hopkins RJ, Hay BA, *et al*. Lung cancer susceptibility model based on age, family history and genetic variants. *PLoS One* 2009;4:e5302.
22 Cohen BH, Diamond EL, Graves CG, *et al*. A common familial component in lung cancer and chronic obstructive pulmonary disease. *Lancet* 1977;2:523–6.
23 Tockman MS*et al*. Airways obstruction and the risk for lung cancer. *Ann Intern Med* 1987;106:512–8.

24 Young RP, Hopkins RJ. How the genetics of lung cancer may overlap with COPD. *Respirology* 2011;16:1047–55.
25 McKay JD, Hung RJ, Han Y, *et al*. Large-Scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet* 2017;49:1126–32.
26 Lassi G, Taylor AE, Timpson NJ, *et al*. The *CHRNA*5-A3-B4 gene cluster and smoking: from discovery to therapeutics. *Trends Neurosci* 2016;39:851–61.
27 Scott A, Lugg ST, Aldridge K, *et al*. Pro-Inflammatory effects of e-cigarette vapour condensate on human alveolar macrophages. *Thorax* 2018;73:1161–9.
28 Tang M-S, Wu X-R, Lee H-W, *et al*. Electronic-cigarette smoke induces lung adenocarcinoma and bladder urothelial hyperplasia in mice. *Proc Natl Acad Sci U S A* 2019;116:21727–31.
29 Young RP, Hopkins RJ. Incorporating genomic data into multivariate risk models for lung cancer. *Genet Med* 2013;15:667–8.
30 Young RP, Hopkins RJ, Duan F, *et al*. Genetic-based approach to stratifying risk of lung cancer outperforms the Brock PLCO2012 model – optimisation of CT screening outcomes in the NLST_ACRIN sub-study (N=10,054). *Am J Repsir Crit Care Med* 2018;197:A4422.
31 Young RP, Hopkins RJ. Chronic obstructive pulmonary disease (COPD) and lung cancer screening. *Transl Lung Cancer Res* 2018;7:347–60.
32 Rivera MP, Tanner NT, Silvestri GA, *et al*. Incorporating coexisting chronic illness into decisions about patient selection for lung cancer screening. An official American thoracic Society research statement. *Am J Respir Crit Care Med* 2018;198:e3–13.
33 VanderWeele TJ, Asomaning K, Tchetgen Tchetgen EJ, *et al*. Genetic variants on 15q25.1, smoking, and lung cancer: an assessment of mediation and interaction. *Am J Epidemiol* 2012;175:1013–20.
34 Improgo MRD, Scofield MD, Tapper AR, *et al*. From smoking to lung cancer: the CHRNA5/A3/B4 connection. *Oncogene* 2010;29:4874–84.
35 Lam DC-L, Luo SY, Fu K-H, *et al*. Nicotinic acetylcholine receptor expression in human airway correlates with lung function. *Am J Physiol Lung Cell Mol Physiol* 2016;310:L232–9.
36 Young RP, Hopkins RJ, Duan F, *et al*. Genetic risk score from 12 SNP panel predicts lung cancer lethality in the National lung screening trial (NLST) – a validation study in the NLST_ACRIN (N=10,054). *Am J Respir Crit Care Med* 2018;197:A2475.
37 DeMeo DL, Mariani T, Bhattacharya S, *et al*. Integration of genomic and genetic approaches implicates *IREB2* as a COPD susceptibility gene. *Am J Hum Genet* 2009;85:493–502.
38 Mazzone PJ, Sears CR, Arenberg DA, *et al*. Evaluating molecular biomarkers for the early detection of lung cancer: when is a biomarker ready for clinical use? an official American thoracic Society policy statement. *Am J Respir Crit Care Med* 2017;196:e15–29.