

Journal of the Korean Society for Research on Nicotine and Tobacco

Manuscript Number: KSRNT22-018

Manuscript Type: 원저

Title: Interaction between Smoking and genetic risk of FEV1/FVC ratio in chronic lung disease: a cross-sectional study

1 **Interaction between Smoking and genetic risk of FEV₁/FVC ratio** 2 **in chronic lung disease: a cross-sectional study**

3 4 **Background**

5 Although both lung function and smoking are risk factors for Chronic Lung Disease (CLD), it is unclear
6 whether the Genetic Risk Score (GRS) of lung function is linked to CLD prevalence and whether
7 smoking has a synergistic effect. We wanted to see if there was an association between GRS for the
8 forced expiratory volume in 1 second divided by the forced vital capacity (FEV₁/FVC) and CLD, and
9 if so, how smoking affected that association.

10 **Methods**

11 After performing a Genome-Wide Association Study (GWAS) of the FEV₁/FVC, we generated
12 FEV₁/FVC-related GRS. Multivariable regression was used to estimate the Odds Ratio (OR) and 95%
13 confidence intervals for the association of GRS for FEV₁/FVC with smoking and CLD.

14 **Results**

15 This study included 8,166 participants (mean of age 52.18 years, 53.45% female). GRS, which
16 contained 62 significant single nucleotide polymorphisms (SNPs) for the FEV₁/FVC ratio across the
17 genome, showed a strong association with CLD. The OR of those in the genetically low lung function
18 and the heavy smoking group was 3.01 times (95% CI 1.07-10.7) than in the genetically high lung
19 function and heavy smoking group. Although there was no significant interaction between GRS for
20 CLD prevalence and smoking status (p for interaction 0.44), the OR for former smokers with genetically
21 low lung function compared to non-smokers with genetically high lung function was 4.29 (95% CI
22 1.20 -17.1).

23 **Conclusions**

24 The lung function of GRS, which was composed of significant SNPs in the GWAS for FEV₁/FVC,
25 demonstrated a significant association with CLD. Participants with genetically low lung function were
26 more likely to develop CLD when exposed to smoking.

27 **Keyword**

28 Chronic lung disease, Cross-sectional study, Smoking, Genetic risk score, Lung function, Interaction

29 만성폐질환에 미치는 FEV₁/FVC 비율의 유전적 위험과 흡연 30 간의 상호작용: 한국인유전체역학조사사업 자료를 이용한 단 31 면 연구 32

33 배경

34 만성폐질환(Chronic Lung Disease, CLD)은 폐와 호흡계에 영향을 미치는 질환이다.
35 만성폐질환의 유형에는 천식, 만성 폐쇄성 폐질환(Chronic Obstructive Pulmonary
36 Disease, COPD), 폐 섬유증, 석면폐증, 폐렴 등이 있다. 이 중 COPD와 천식은 만성폐
37 질환의 가장 흔한 유형으로 2017년에는 전 세계적으로 320만 명이 COPD로 사망했으
38 며, 이는 1990년의 사망자 수보다 23% 더 많았다 [1].

39 흡연은 만성 염증 반응과 산화 스트레스를 주로 일으키는 만성폐질환의 주요 위험
40 이다. 이전 연구에서 흡연을 줄이면 COPD의 위험을 줄일 수 있다고 보고되었다[2, 3].
41 그러나 COPD 발생의 25-45%는 비흡연자에서 발생하였고 흡연을 지속하는 대상자 중
42 25%에서 COPD가 발생했다 [4, 5]. 이는 만성폐질환 발생에 다른 위험요인이 있음을
43 나타내며, 다른 위험요인 중 하나는 폐기능의 유전적 요소이다.

44 지난 10년 동안 전장 유전체 연관 분석(Genome-Wide Association Studies,
45 GWAS)은 유전자 TNS1(2q35), GSTCD(4q24), HTR4(5q33), AGER(6p21) 및
46 THSD4(15q23)를 포함하여 폐기능 지표인 노력성 폐활량(Forced Vital Capacity,
47 FVC), 1초간 노력성 호기량(Forced Expiratory Volume, FEV₁), FEV₁/FVC 비율과 관
48 련된 여러 유전적 변이를 확인하였다 [6-10]. 그러나 폐기능과 일관되고 유의미하게 관
49 련된 각각의 유전적 변이의 영향은 작았으며, 이는 개별 유전 변이가 폐기능의 일-부분
50 만을 설명한다는 것을 보여주었다. 이를 해결하기 위한 방법으로 작은 영향력을 가진 여
51 러 개의 단일 염기 다형성(Single Nucleotide Polymorphism, SNP)을 합산한 유전자 위
52 험 점수(Genetic Risk Score, GRS)를 통해서 복잡한 질병의 유전적 위험을 밝힐 수 있
53 었다 [11]. 최근 환자-대조군 연구를 기반으로, 작은 영향력을 가진 수백만 개의 단일
54 염기 다형성을 포함하는 유전자 위험 점수가 COPD를 비롯한 다양한 질병 발생 위험을
55 높이는 것을 예측하였다[12, 13]. 그러나 폐기능 관련 유전자 위험 점수가 만성폐질환과
56 관련이 있는지 여부와 흡연 및 폐기능 관련 유전적 요인이 만성폐질환에 시너지 효과를
57 주는지 여부는 여전히 논란의 여지가 있다.

58 따라서, 이 연구에서는 폐기능 관련 유전적 위험이 만성폐질환 위험과 관련되는지

59 여부와 이 연관성이 흡연의 영향에 따라 차이가 있는지 여부를 평가하였다.

60

61 방법

62 1. 연구 자료 및 연구 대상

63 이 연구 자료는 2001년부터 2002년까지 안산(도시)과 안성(농촌)에서 한국인유전
64 체역학조사사업(Korean Genome and Epidemiology Study, KoGES)에 참여한 40세
65 에서 69세까지 총 10,030명이었다. [14]. 분석에는 성별, 나이, 흡연 상태와 폐기능,
66 만성폐질환 여부 등 주요 분석 변수에 대한 정보가 누락된 대상자를 제외하였고, 최종
67 8,166(남성 3,826명, 여성 4,340명)명이 분석에 포함되었다. 이 연구는 세브란스병원
68 연구심의위원회의 승인을 받았다(과제승인번호 4-2015-1132). 연구대상자의 음주
69 상태는 음주 양상에 따라 비음주, 과거 음주, 현재 음주로 분류하였고, 소득 관련 변수
70 는 월 소득 100만원을 기준으로 저소득과 고소득으로 구분하였으며, 교육수준은 최종
71 학력이 고등학교 이하인 군을 낮음, 전문대 이상인 군을 높음으로 분류하였다.

72 2. 유전 위험 점수(Genetic Risk Score, GRS) 산출

73 유전 위험 점수는 일반적으로 사용되고 있는 가중치(Weighted) 유전 위험 점수
74 방법으로 계산되었다. 이것은 대규모 전장 유전체 연관 분석에서 추정된 특정 형질과
75 관련된 유전자 형질의 유전적 효과를 회귀 계수 값으로 하여 가중치를 주는 유전 위
76 험 점수 계산 방법이다. 유전 위험 점수 산출법은 각 개인이 가지고 있는 선발된 단
77 일 염기 다형성의 위험 형질(risk allele)의 수(0, 1, 2)를 전체 집단에 대해 가산적
78 (additive) 모형을 가정하여 유전적 형질의 증가 당 회귀 계수 값을 구한 후 위험 형
79 질의 수, 즉, 0, 1, 2에 가중치로서 회귀 계수를 곱한 후 더하는 방식이다. 이 연구에
80 서 폐기능 관련 유전 위험 점수는 삼분위수에 따라 높음, 중간, 낮음으로 나누어 분
81 류하였다.

$$82 \quad PRS = \sum_{i=1}^n \text{Number of risky allele in } SNP_i \times \text{Weight}_i$$

83 n : 전체 대상자수, Weight_i : i 번째 대상자의 가중치, SNP_i : 각각의 SNP

84

85 3. 흡연 상태 및 흡연량

86 연구 참여자의 흡연 양상은 흡연 상태에 따라 비흡연, 과거 흡연, 현재 흡연으로
87 분류하였다. 평생 흡연량은 일일 평균 흡연량에 흡연 기간을 곱하는 방법으로 현재

88 흡연자와 과거흡연자의 갑년(pack year, PY) 변수를 생성하였다. 갑년의 계산방법은
89 하루 흡연량(갑) x곱하기 흡연 기간(년)이다. 평생 흡연량에 따라 비흡연(0), 가벼운
90 흡연(0.1-14.9개비), 중등도 흡연(15-27.9개비), 심한 흡연(≥ 28 개비)으로 분류
91 하였다.

92 4. 만성폐질환 및 폐기능

93 폐기능은 기준선에서 미국의 캘리포니아주 소재 Sensormedics Corporation의
94 VMAX2130 폐활량계로 측정된 자료를 사용했다. 만성폐질환은 FEV₁/FVC 비율이
95 0.70 미만인 경우와 기준선에서 보고된 자가설문자료를 이용하였다.

96 5. 통계 분석

97 인구사회학적 특성, 흡연여부, 흡연량, 폐기능 관련 유전위험점수의 분포와
98 기술통계를 제시하였으며, 성별에 따른 인구학적 변인 차이를 알아보기 위해
99 Student t-검정과 카이제곱 검정(Chi-square test)을 사용하였다. 연령, 성별 보정
100 및 연령, 성별, 교육수준, 사회경제적 지위, 체질량 지수, 음주양상 보정 두 가지 모
101 형의 단순회귀분석으로 만성폐질환 관련요인을 분석하였다. 폐기능 관련 유전 점수
102 를 산출하기 위해 전장 유전체 연관 분석을 실시하였으며, 다중분석보정을 위해 P-
103 value임계값 5×10^{-5} 에 따라 유의한 단일 염기 다형성을 선별하였다. 선별된 단일
104 염기 다형성은 연관 불평형(linkage disequilibrium, LD)분석을 수행하여 독립적인
105 단일 염기 다형성을 유전 위험 점수를 산출하는데 사용하였다. 최종적으로 만성폐질
106 환 위험에 대한 흡연과 폐기능 관련 유전 점수와의 상호작용 검정을 수행하였다. 상
107 호작용 효과의 대한 통계적 유의성은 상호작용항의 F변화량 유의확률로 $P < 0.05$ 를
108 기준으로 하였다. 모든 분석에는 R 버전 4.2.1(R Center for Statistical Computing,
109 Vienna, Austria)과 Plink 프로그램 version 1.09을 이용하였으며, 통계적 유의성
110 검정은 $P < 0.05$ (양측 검정)로 하였다.

111

112 결과

113 이 연구의 8,166명의 평균 연령은 52.18세이었고, 53.45%가 여성이었다(표1). 연
114 구 대상자 중 과거 흡연자는 1,252명(15.3%), 현재 흡연자는 2,037명(24.9%)이었으
115 며, 이 중 1,048명은 중등도 흡연(15-27.9갑년), 1,037명은 심각한 흡연(≥ 28 갑년)의
116 경험이 있었다. 만성폐질환 발생 수는 총 50명(0.6%)이었다.

117 FEV₁/FVC 비율에 대해 62개의 중요한 단일 염기 다형성을 포함하는 폐기능 유전
118 위험 점수는 만성폐질환과 강한 연관성을 보여주었다(그림1, 2, 보충 표 3). 폐기능 관련
119 낮은 유전 점수와 비교하여 높은 폐기능 유전 점수는 만성폐질환과 유의미한 연관 관계
120 를 보였다(odds ratio (OR)=2.57, 95% confidence interval (CI)=1.27~5.61). 또한 폐
121 기능에 대한 유전적 위험 점수가 낮을수록 만성폐질환 유병률이 증가하는 경향을 보였다
122 (P for trend=0.008)(표 2).

123 만성폐질환의 유병률은 흡연 상태와 흡연량(pack-years)에 따라 변화했다. 과거
124 흡연과 현재 흡연의 만성폐질환에 대한 발생 위험은 비흡연 그룹과 비교하여 OR은 각각
125 1.92(95% CI =0.73~5.27) 및 1.59(95% CI =0.64~4.16)이었다 (표 3). 또한, 가벼운
126 흡연 대비 중등도 흡연, 심각한 수준의 흡연과 만성폐질환 발생 위험 사이 OR은 각각
127 1.31(95% CI =0.41~4.54) 및 2.32(95% CI =0.87~7.44)이었다 (표 3).

128 폐기능 관련 유전 점수와 흡연을 결합하여 전체 대상자를 그룹화한 결과, 만성폐질
129 환 위험에 미치는 흡연과 폐기능 관련 유전 점수 사이 유의한 상호작용은 없었지만 만성
130 폐질환 유병률은 흡연과 유전 점수에 따라 증가했다(p for interaction=0.44)(표 4). 폐
131 기능 관련 유전 점수가 높고 비흡연자인 대상자에 비해, 폐기능 관련 유전 점수가 낮고
132 과거 흡연자인 대상자의 OR은 4.29(95% CI =1.20~17.1)이었다. 비교적 유전적으로
133 폐기능이 낮은 현재 흡연자의 만성폐질환 발생 위험은 2.97%로 낮았다 (표4). 폐기능
134 유전 위험 점수와 흡연량 그룹에서 유사한 패턴이 관찰되었으며, 유전 점수가 낮고 흡연
135 량이 많은 그룹에서 가장 높은 위험이 관찰되었지만 이는 통계적으로 유의하지 않았다
136 (OR=2.82, 95% CI =0.86~12.7) (표 4). 만성폐질환에 미치는 흡연량에 따른 양의 연
137 관성은 유전적 점수가 낮은 그룹에서만 관찰되었지만 이는 통계적으로 유의하지 않았다
138 (OR=2.15, 95% CI =0.78~7.00)(보충표 1). 폐기능 유전 위험 점수를 이분화하여 분
139 석한 결과, 높은 유전 점수를 가진 심각한 수준의 흡연을 하는 대상자 대비 낮은 유전
140 점수를 가진 심각한 수준의 흡연을 하는 대상자의 만성폐질환의 오즈비는 3.01배(95%
141 CI =1.07~10.7) 높았다 (보충표 2).

142

143 고찰

144 이 연구는 지역사회 일반인구를 대상으로 분석한 결과 FEV₁/FVC에 대한 전장 유전
145 체 연관 분석의 중요한 단일 염기 다형성으로 구성된 폐기능 유전 위험 점수는 만성폐질
146 환과의 높은 인과적 연관성을 보여주었다. 이는 폐기능에 유전적 영향이 크다는 것을 보

147 여주며, 유전 위험 점수가 만성폐질환 발생의 예측력을 크게 향상시킬 수 있음을 보여준
148 다. 이전 다른 연구에서도 유전 위험 점수가 폐기능뿐만 아니라 만성폐질환의 주요 특징
149 인 폐기능 저하와도 관련이 있음을 보여주었다 [15, 16]. 따라서 이 연구는 폐기능 유전
150 위험 점수가 만성폐질환의 예측 도구로 적용될 수 있음을 보여주며, 대규모 전향적 코호
151 트 연구의 근거를 제시하였다.

152 만성폐질환에 대한 유전 위험 점수 연구는 동양인을 대상으로 한 연구에서 찾아보기
153 어려웠다. 이에 대한 주된 이유는 대부분의 전장 유전체 연관 분석은 서양인에서 주로
154 연구되기 때문이다. 따라서 한국인을 대상으로 한 유전 위험 점수 분석은 한국인 특성
155 정밀 의학 및 예방을 구현할 수 있다.

156 흡연은 산화 스트레스를 통해 조직 손상을 직접 유발하거나 간접적으로 염증반응을
157 유도한다. 흡연과 폐기능 관련 유전적 감수성이 만성폐질환 발생, 발달 및 폐기능 저하
158 와 상호작용하는지 여부는 오랫동안 논의되었다. 이 연구는 폐기능 관련 유전 점수가 낮
159 은 집단에서 흡연이 만성폐질환 유병에 중요한 영향을 미친다는 것을 확인한 한국인 기
160 반 첫 번째 연구이다. 다른 인종을 대상으로 한 유사한 연구로 낮은 FEV₁/FVC(26개의
161 단일 염기 다형성으로 구성된 유전 위험 점수) 비율에 대한 유전적 위험이 높은 대상자
162 는 흡연의 유해한 영향에 더 취약했다 [17]. 15q25의 CHRNA3 및 CHRNA5 를 포함
163 한 여러 유전자는 폐기능, COPD 및 흡연 행동과 밀접한 관련이 있었다 [18–20]. 폐기
164 능 관련 높은 유전 위험 점수는 다른 흡연 습관 및 다양한 니코틴 함량을 가진 담배 선
165 택을 포함하여 측정되지 않은 흡연 노출로 이어질 수 있다. 그러나 UK Biobank의 연구
166 에 따르면 흡연과 폐기능의 유전적 영향은 독립적으로 작용한다. 이 연구에서 발견되지
167 않은 상호작용은 다른 연구에서 찾아볼 수 있었다. 약 5만명을 대상으로 FEV₁에 대한
168 전장 유전체 연관 분석 결과를 이용한 연구결과는 심각한 흡연자와 비흡연자 사이 FEV₁
169 의 유전적 감수성이 낮아 흡연과 상호작용하는 폐기능 관련 유전적 영향은 일부분임을
170 보여주었다[21]. 또한, 유전 위험 점수를 구성하기 위해 279개의 FEV₁, FVC 관련 단일
171 염기 다형성을 결합한 단면 연구 결과에 따르면, FEV₁에 대한 흡연-유전 위험 점수의
172 상호작용은 없었으며, COPD에 대한 FVC와 흡연 사이 약한 상호작용만 관찰되었다 [8].
173 5,070명의 참가자로 구성된 전장 유전체 연관 분석 또한 5개의 FEV₁/FVC 관련 유전
174 변이는 발견했지만, 이들과 흡연 사이의 상호작용은 발견되지 않았다 [20].

175 폐기능 관련 유전적 위험이 높은 대상자는 만성폐질환에 대비할 수 있다. 유전 점수
176 에 따라 판별된 고위험군은 건강한 생활 방식을 선택할 수 있으며, 이는 $\alpha 1-$

177 antitrypsin 결핍 연구에서 확인되었다 [22]. 유전 위험 점수 정보를 바탕으로한 조기
178 개입이 만성폐질환 진단을 개선하고, 중증 만성폐질환의 부담을 줄일 수 있는지 여부는
179 더 심도 있는 연구가 필요하다.

180 이 연구는 한국인을 대상으로 폐기능에 대한 전장 유전체 연관 분석을 한 뒤 폐기능
181 관련 유전 위험 점수를 생성하였고, 이를 이용하여 흡연과 폐기능 관련 유전 위험 점수
182 와의 상호작용이 만성폐질환에 미치는지 확인했다는 강점이 있다.

183 이 연구의 제한점은 다음과 같다. 첫째, 적은 사례 수로 인해 설명력이 낮을 수 있
184 다. 둘째, 흡연 관련 변수는 자가설문자료로 회상편향과 정확도 저하가 나타날 수 있다.
185 셋째, 희귀 유전 변이가 본 연구에 포함되지 않았기 때문에 유전적 위험도 측정이 부정
186 확할 수 있다. 넷째, 이 연구는 한국인을 대상으로 한 연구로 결과를 다른 인종에 일반
187 화하기 어렵다. 다섯째, 본 연구결과는 만성폐질환에 이환된 대상자들이 금연하여 나타
188 나는 단면연구로 인한 측정오류가 포함될 수 있다. 마지막으로 이 연구에서는 질병의 중
189 중도에 대한 정보를 반영하지 못해 잠재적인 용량-반응 관계를 놓칠 수 있다.

190

191 결론

192 인구 기반 연구에서 폐기능과 연관성이 있는 62개의 단일 염기 다형성으로 구성된
193 폐기능 유전 위험 점수는 만성폐질환과 유의미한 연관성을 보여주었다. 또한 이 연구는
194 폐기능 관련 유전 점수가 낮은 대상자가 흡연에 노출될 경우 만성폐질환 발생에 더 취약
195 할 수 있음을 발견했다.

196 **요약**

197 **배경:** 폐기능과 흡연 모두 만성폐질환의 위험요소이지만 폐기능 관련 유전 위험 점수가
198 만성폐질환 유병률과 연관되어 있는지를 확인한 연구는 충분치 않으며, 유전적 위험 점
199 수와 흡연 사이 상호작용의 유무는 불분명하다. 이에 본 연구에서는 폐기능 지표 중 하
200 나인 FEV₁/FVC의 유전 점수가 만성폐질환에 미치는 영향을 확인하고, 나아가 흡연이
201 이 연관성에 어떤 영향을 미치는지 확인하고자 한다.

202 **방법:** FEV₁/FVC의 전장 유전체 연관 분석을 수행한 후 FEV₁/FVC 관련 유전 변이들로
203 유전 위험 점수를 생성했다. FEV₁/FVC에 대한 유전 위험 점수와 흡연이 만성폐질환에
204 미치는 영향에 대한 오즈비(OR) 및 95% 신뢰 구간을 추정하기 위해 다변수 로지스틱
205 회귀분석을 사용했다.

206 **결과:** 이 연구에는 8,166명의 대상자가 포함되었다(평균 연령 52.18세, 여성 53.45%).
207 FEV₁/FVC 비율에 대해 62개의 중요한 단일염기다형성을 포함하는 유전 위험 점수는
208 만성폐질환과 강한 연관성이 있었다. 유전적으로 높은 폐기능과 비교하여 유전적으로 낮
209 은 폐기능의 OR은 2.57(95% 신뢰구간 1.27-5.61)이었다. 유전적으로 폐기능이 낮으
210 면서 심한 수준의 흡연을 하는 대상자는 유전적으로 폐기능이 높으면서 심각한 수준의
211 흡연을 하는 사람보다 만성폐질환의 OR이 3.01배(95% 신뢰구간 1.07-10.7) 높았다.
212 만성폐질환에 대한 유전 위험 점수와 흡연 상태 사이에 유의한 상호작용은 없었지만(상
213 호작용에 대한 P=0.44), 유전적으로 폐기능이 높은 비흡연자와 비교하여 유전적으로 낮
214 은 폐기능을 가진 과거 흡연자의 만성폐질환의 OR은 4.29배 높았다(95% 신뢰구간
215 1.20 -17.1).

216 **결론:** FEV₁/FVC에 대한 전장 유전체 연관 분석의 중요한 단일염기다형성으로 구성된
217 유전 위험 점수에 따른 폐기능은 만성폐질환과 높은 연관성을 보여주었다. 유전적으로
218 폐기능이 낮은 대상자는 흡연에 노출되었을 때 만성폐질환이 발생할 가능성이 더 높았다.

219 **중심단어:** 만성폐질환, 흡연, 유전적 위험 점수, 폐기능, 상호작용

참고문헌

1. Li X, Cao X, Guo M, Xie M, Liu X. Trends and risk factors of mortality and disability adjusted life years for chronic respiratory diseases from 1990 to 2017: systematic analysis for the Global Burden of Disease Study 2017. *BMJ* 2020; 368.
2. Celli BR, Decramer M, Wedzicha JA, Wilson KC, Agustí A, Criner G J, et al. An official American Thoracic Society/European Respiratory Society statement: research questions in chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 2015; 191(7), e4–e27.
3. Postma DS, Bush A, van den Berge M. Risk factors and early origins of chronic obstructive pulmonary disease. *The Lancet* 2015; 385(9971), 899–909.
4. Løkke A, Lange P, Scharling H, Fabricius P, Vestbo J. Developing COPD: a 25 year follow up study of the general population. *Thorax* 2006; 61(11), 935–939.
5. Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *The Lancet* 2009; 374(9691), 733–743.
6. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat ME, et al. Genome-wide association study identifies five loci associated with lung function. *Nature Genetics* 2010; 42(1), 36–44.
7. Loth DW, Artigas MS, Gharib SA, Wain LV, Franceschini N, Koch B, et al. Genome-wide association analysis identifies six new loci associated with forced vital capacity. *Nature Genetics* 2014; 46(7), 669–677.
8. van der Plaat DA, de Jong K, Lahousse L, Faiz A, Vonk JM, van Diemen CC, et al. Genome-wide association study on the FEV1/FVC ratio in never-smokers identifies HHIP and FAM13A. *Journal of Allergy and Clinical Immunology* 2017; 139(2), 533–540.
9. Yao TC, Du G, Han L, Sun Y, Hu D, Yang JJ, et al. Genome-wide association study of lung function phenotypes in a founder population. *Journal of allergy and clinical immunology* 2014; 133(1), 248–255.
10. Lutz SM, Cho MH, Young K, Hersh CP, Castaldi PJ, McDonald ML, et al. A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC genetics* 2015; 16(1), 1–11.
11. Igo Jr RP, Kinzy TG, Cooke Bailey JN. Genetic risk scores. *Current protocols in human genetics* 2019; 104(1), e95.

12. Moll M, Sakornsakolpat P, Shrine N, Hobbs BD, DeMeo DL, John, C, et al. Chronic obstructive pulmonary disease and related phenotypes: polygenic risk scores in population-based and case-control cohorts. *The Lancet Respiratory Medicine* 2020; 8(7), 696–708.
13. Elliott J, Bodinier B, Bond TA, Chadeau-Hyam M, Evangelou E, Moons KG, et al. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *Jama* 2020; 323(7), 636–645.
14. Kim Y, Han BG, KoGES Group. Cohort profile: the Korean genome and epidemiology study (KoGES) consortium. *International journal of epidemiology* 2017; 46(2), e20–e20.
15. John C, Artigas MS, Hui J, Nielsen SF, Rafaels N, Paré PD, et al. Genetic variants affecting cross-sectional lung function in adults show little or no effect on longitudinal lung function decline. *Thorax* 2017; 72(5), 400–408.
16. Oelsner EC, Ortega VE, Smith BM, Nguyen JN, Manichaikul AW, Hoffman EA, et al. A genetic risk score associated with chronic obstructive pulmonary disease susceptibility and lung structure on computed tomography. *American journal of respiratory and critical care medicine* 2019; 200(6), 721–731.
17. Aschard H, Tobin MD, Hancock DB, Skurnik D, Sood A, James A, et al. Evidence for large-scale gene-by-smoking interaction effects on pulmonary function. *International journal of epidemiology* 2017; 46(3), 894–904.
18. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Artigas MS, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *The Lancet Respiratory Medicine* 2015; 3(10), 769–781.
19. Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nature genetics* 2010; 42(5), 436–440.
20. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature genetics*, 2010; 42.5: 441–447.
21. Shrine N, Guyatt AL, Erzurumluoglu AM, Jackson VE, Hobbs BD, Melbourne CA, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nature genetics* 2019; 51(3), 481–493.

22. Carpenter MJ, Strange C, Jones Y, Dickson MR, Carter C, Moseley MA, et al. Does genetic testing result in behavioral health change? Changes in smoking behavior following testing for alpha-1 antitrypsin deficiency. *Annals of behavioral medicine* 2007; 33(1), 22–28.

Table 1. Baseline characteristics of the study population (n = 8,166)

Variable	Men (n=3,826)	Women (n=4,340)	P-value
Age, yr	51.8 (8.85)	52.5 (9.00)	<0.001
Body mass index, kg/m ²	24.2 (2.92)	24.9 (3.25)	<0.001
Smoking status			<0.001
- Never	733 (19.2%)	4048 (95.1%)	
- Former	1195 (31.3%)	57 (1.34%)	
- Current	1885 (49.4%)	152 (3.57%)	
Smoking pack-years			<0.001
- No (0)	867 (22.7%)	4156 (95.8%)	
- Light (0.1-14.9)	904 (30.6%)	144 (78.3%)	
- Intermediate (15-27.9)	1023 (34.6%)	25 (13.6%)	
- Heavy (≥28)	1032 (34.9%)	15 (8.15%)	
Drinking status			<0.001
- Never	690 (18.1%)	3053 (71.1%)	
- Former	367 (9.63%)	131 (3.05%)	
- Current	2754 (72.3%)	1110 (25.9%)	
Low household income (> 1 million won)	1036 (27.1%)	1761 (40.6%)	<0.001
Education			<0.001
- Lower qualification	2980 (78.3%)	4041 (94.0%)	
- Higher qualification	828 (21.7%)	257 (5.98%)	
Genetic risk category			0.741
- Low (lowest tertile)	1279 (33.4%)	1439 (33.2%)	
- Intermediate	1261 (33.0%)	1465 (33.8%)	
- High (highest tertile)	1286 (33.6%)	1436 (33.1%)	
FVC	4.27 (0.70)	3.06 (0.63)	<0.001
FEV ₁	3.32 (0.64)	2.49 (0.48)	<0.001
FEV ₁ /FVC (%)	77.9 (8.44)	81.7 (6.68)	<0.001
HDL-C, mg/dL	44.5 (9.51)	46.3 (9.71)	<0.001
Triglycerides, mg/dL	163 (80.9)	142 (68.3)	<0.001
Chronic Lung Disease	33 (0.86%)	17 (0.39%)	0.01

Table 2. Risk of chronic lung disease (CLD) according to genetic risk score of lung function (n=8,166)

Genetic risk score of lung function	Total participants, n	CLD cases, n (%)	Age, sex-adjusted Model		Fully-adjusted Model*	
			Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value
High	2718	26 (0.95)	1.00 (reference)		1.00 (reference)	
Intermediate	2726	14 (0.51)	1.44 (0.64-3.35)	0.380	1.46 (0.65-3.39)	0.37
Low	2722	10 (0.36)	2.58 (1.28-5.63)	0.011	2.57 (1.27-5.61)	0.012
p for trend				0.008		0.008

*Fully-adjusted models were adjusted for age, sex, education, socioeconomic status (household income), body mass index, and alcohol consumption.

†CI = Confidence Interval

Table 3. Risk of chronic lung disease (CLD) according to smoking categories

Smoking subgroup	Total participants, n	CLD cases, n (%)	Age, sex-adjusted Model		Fully-adjusted Model*	
			Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value
Smoking status						
- Never smoking	4781	19 (0.39)	1.00 (reference)		1.00 (reference)	
- Former smoking	1252	13 (1.03)	2.04 (0.77-5.60)	0.160	1.92 (0.73-5.27)	0.200
- Current smoking	2037	18 (0.88)	1.92 (0.77-4.95)	0.170	1.59 (0.64-4.16)	0.330
- p for trend				0.250		0.480
Smoking pack-years						
- Light (0.1-14.9) smoking	1048	5 (0.47)	1.00 (reference)		1.00 (reference)	
- Intermediate (15-27.9) smoking	1048	7 (0.66)	1.43 (0.45-4.92)	0.550	1.31 (0.41-4.54)	0.660
- Heavy (≥ 28) smoking	1047	17 (1.62)	2.39 (0.90-7.63)	0.100	2.32 (0.87-7.44)	0.120
- p for trend				0.083		0.088

*Fully-adjusted models were adjusted for age, sex, education, socioeconomic status (household income), body mass index, and alcohol consumption.

†CI = Confidence Interval

Table 4. Risk of chronic lung disease (CLD) according to a) genetic risk and smoking status or b) genetic risk and smoking pack-years.

a) Genetic risk score (GRS) of lung function and smoking status (n=8,070)

Subgroup	Total participants, n	CPD cases, n (%)	Age, sex-adjusted Model		Fully-adjusted Model		P for interaction
			Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value	
High genetic score							
- Never smoking	1610	5 (0.31)	1.00 (reference)		1.00 (reference)		0.44
- Former smoking	419	2 (0.47)	1.29 (0.17-6.98)	0.78	1.22 (0.16-6.61)	0.82	
- Current smoking	653	3 (0.45)	1.35 (0.25-6.43)	0.71	1.11 (0.20-5.32)	0.9	
Intermediate genetic score							
- Never smoking	1606	5 (0.31)	1.06 (0.29-3.81)	0.93	1.06 (0.29-3.82)	0.93	
- Former smoking	408	3 (0.73)	1.96 (0.36-9.51)	0.41	1.90 (0.35-9.20)	0.43	
- Current smoking	682	6 (0.87)	2.63 (0.68-10.8)	0.17	2.21 (0.57-9.15)	0.26	
Low genetic score							
- Never smoking	1565	9 (0.57)	2.01 (0.69-6.55)	0.21	2.01 (0.69-6.58)	0.21	
- Former smoking	425	8 (1.88)	4.74 (1.31-19.0)	0.021	4.29 (1.20-17.1)	0.03	
- Current smoking	702	9 (1.28)	3.56 (1.04-13.7)	0.051	2.97 (0.86-11.5)	0.1	

b) Genetic risk score (GRS) of lung function and smoking pack-years (n=3,143)

Subgroup	Total participants, n	CPD cases, n (%)	Age, sex-adjusted Model		Fully-adjusted Model		P for interaction
			Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value	
High genetic score							
- Light (0.01-19.99) smoking	535	3 (0.56)	1.00 (reference)		1.00 (reference)		0.52
- Heavy (≥ 20.00) smoking	498	2 (0.41)	0.57 (0.07-3.51)	0.55	0.57 (0.07-3.48)	0.54	
Intermediate genetic score							
- Light (0.01-19.99) smoking	505	4 (0.79)	1.49 (0.33-7.63)	0.6	1.55 (0.34-7.97)	0.57	
- Heavy (≥ 20.00) smoking	534	4 (0.74)	1.06 (0.23-5.48)	0.94	1.09 (0.23-5.62)	0.92	
Low genetic score							
- Light (0.01-19.99) smoking	532	4 (0.75)	1.31 (0.29-6.68)	0.73	1.26 (0.27-6.45)	0.77	
- Heavy (≥ 20.00) smoking	1571	18 (1.14)	2.94 (0.91-13.2)	0.1	2.82 (0.86-12.7)	0.12	

*Fully-adjusted models were adjusted for age, sex, education, socioeconomic status (household income), body mass index, and alcohol consumption.

†CI = Confidence Interval

Supplementary Table 1. Risk of chronic lung disease (CLD) according to a smoking pack-year category within each genetic risk of lung function category

Subgroup	Total participants, n	CPD cases, n (%)	Age, sex-adjusted Model		Fully-adjusted Model	
			Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value
High genetic score (≥ 0)						
- Light (0.01-19.99) smoking	805	6 (0.74)	1.00 (reference)		1.00 (reference)	
- Heavy (≥ 20.00) smoking	759	4 (0.52)	0.51 (0.13-1.84)	0.31	0.53 (0.13-1.92)	0.34
Low genetic score (< 0)						
- Light (0.01-19.99) smoking	767	5 (0.65)	1.00 (reference)		1.00 (reference)	
- Heavy (≥ 20.00) smoking	812	14 (1.72)	2.13 (0.78-6.89)	0.17	2.15 (0.78-7.00)	0.16

*Fully-adjusted models were adjusted for age, sex, education, socioeconomic status (household income), body mass index, and alcohol consumption.

†CI = Confidence Interval

Supplementary Table 2. Risk of chronic lung disease (CLD) according to binary genetic risk and smoking pack-years.

Subgroup	Total participants, n	CPD cases, n (%)	Age, sex-adjusted Model		Fully-adjusted Model	
			Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value
High genetic score (≥ 0)						
- Light (0.01-19.99) smoking	805	6 (0.74)	1.82 (0.51-7.25)	0.36	1.83 (0.51-7.32)	0.36
- Heavy (≥ 20.00) smoking	759	4 (0.52)	1.00 (reference)		1.00 (reference)	
Low genetic score (< 0)						
- Light (0.01-19.99) smoking	767	5 (0.65)	1.54 (0.40-6.32)	0.52	1.54 (0.40-6.33)	0.53
- Heavy (≥ 20.00) smoking	812	14 (1.72)	3.11 (1.11-11.0)	0.04	3.01 (1.07-10.7)	0.05

*Fully-adjusted models were adjusted for age, sex, education, socioeconomic status (household income), body mass index, and alcohol consumption.

†CI = Confidence Interval

Supplementary Table 3. SNPs associated with the FEV₁/FVC ratio in the identification analysis.

CHR	SNP	A1	BETA	SE	L95	U95	P
1	rs56758642	A	0.4684	0.1103	0.2523	0.6846	2.18E-05
1	rs77585818	C	-2.421	0.509	-3.418	-1.423	2.01E-06
1	rs4915931	A	1.247	0.2998	0.659	1.834	3.23E-05
1	rs4949830	T	0.8123	0.1908	0.4384	1.186	2.09E-05
1	rs17884381	A	-1.096	0.2668	-1.619	-0.5729	4.05E-05
1	rs79829191	A	-0.8589	0.2075	-1.266	-0.4522	3.52E-05
1	rs60107954	C	-0.515	0.1266	-0.7631	-0.2669	4.77E-05
2	rs56017407	A	-1.024	0.2263	-1.468	-0.5809	6.05E-06
2	rs1811387	G	1.203	0.2872	0.6402	1.766	2.82E-05
2	rs117966635	A	1.101	0.2711	0.5697	1.633	4.93E-05
2	rs76368657	T	-1.217	0.2976	-1.8	-0.6337	4.37E-05
2	rs4973510	A	-0.4535	0.1093	-0.6677	-0.2393	3.37E-05
3	rs9873658	C	-0.6138	0.1425	-0.8931	-0.3345	1.67E-05
3	rs7622512	C	0.45	0.1086	0.2373	0.6628	3.42E-05
4	rs2972011	A	0.4585	0.108	0.2469	0.6701	2.19E-05
4	rs60955950	C	-0.5551	0.1086	-0.7679	-0.3422	3.26E-07
4	rs58899027	A	-0.5782	0.1422	-0.857	-0.2994	4.84E-05
5	rs59192924	G	-1.323	0.2765	-1.865	-0.7811	1.74E-06
5	rs142751422	G	1.21	0.283	0.6555	1.765	1.92E-05
5	rs73795078	G	-1.06	0.2487	-1.547	-0.5723	2.05E-05
6	rs78979324	T	1.244	0.2827	0.6894	1.798	1.10E-05
6	rs9466788	A	-1.15	0.2676	-1.674	-0.6253	1.75E-05
6	rs386480	C	0.5189	0.1216	0.2805	0.7572	2.00E-05
6	rs34065257	A	-1.089	0.2434	-1.566	-0.6122	7.74E-06
6	rs140109873	C	-1.466	0.3542	-2.161	-0.7723	3.50E-05
6	rs147105918	T	-1.527	0.3562	-2.225	-0.8289	1.83E-05
6	rs12664268	C	0.9981	0.2404	0.5269	1.469	3.33E-05
7	rs2057901	T	-0.6217	0.1389	-0.8939	-0.3495	7.70E-06
7	rs9791769	T	1.072	0.2448	0.5923	1.552	1.21E-05
7	rs150265932	T	-1.443	0.3479	-2.125	-0.7613	3.38E-05
7	rs12234835	A	-0.9966	0.2442	-1.475	-0.5179	4.53E-05
7	rs71542364	C	1.442	0.3474	0.7611	2.123	3.34E-05
7	rs17138464	G	0.5781	0.1418	0.3001	0.856	4.62E-05
8	rs145128976	A	0.789	0.1921	0.4125	1.165	4.03E-05
8	rs114811959	C	1.131	0.2695	0.6025	1.659	2.75E-05
8	rs4871730	C	0.5008	0.1093	0.2865	0.7151	4.70E-06
8	rs7812357	T	0.923	0.2246	0.4827	1.363	4.01E-05
9	rs1935314	C	0.7479	0.1825	0.3902	1.106	4.21E-05

9	rs7034762	T	0.6773	0.1643	0.3552	0.9994	3.80E-05
10	rs117366124	T	1.993	0.487	1.039	2.948	4.30E-05
10	rs143636873	G	1.068	0.2596	0.5596	1.577	3.90E-05
10	rs2815401	C	-0.7972	0.1629	-1.116	-0.4779	1.01E-06
10	rs10829878	C	0.4537	0.1112	0.2357	0.6717	4.56E-05
11	rs78819781	A	-2.226	0.5308	-3.267	-1.186	2.76E-05
11	rs9705224	C	-0.6253	0.1527	-0.9247	-0.326	4.28E-05
11	rs77130158	T	-1.076	0.2598	-1.585	-0.5664	3.50E-05
12	rs74104371	G	-0.6686	0.1635	-0.9891	-0.3481	4.37E-05
12	rs75617170	G	-0.7426	0.1742	-1.084	-0.4012	2.03E-05
12	rs61177363	G	-1.805	0.4355	-2.659	-0.9513	3.44E-05
13	rs1323562	T	0.4607	0.113	0.2393	0.6822	4.60E-05
13	rs9521144	G	-0.4773	0.1094	-0.6916	-0.2629	1.30E-05
14	rs79858042	C	-0.927	0.2281	-1.374	-0.4799	4.87E-05
15	rs899462	G	0.5102	0.1132	0.2884	0.7321	6.65E-06
16	rs12598420	T	-0.5924	0.1404	-0.8677	-0.3172	2.48E-05
17	rs77148002	T	-1.089	0.2614	-1.601	-0.5769	3.12E-05
17	rs2001868	C	0.6112	0.1497	0.3177	0.9046	4.51E-05
18	rs78947871	T	0.898	0.2087	0.4889	1.307	1.71E-05
18	rs60932948	A	-0.6424	0.1469	-0.9302	-0.3545	1.23E-05
20	rs11699819	C	-0.5732	0.1328	-0.8336	-0.3128	1.62E-05
20	rs2206414	A	-0.4781	0.1168	-0.7071	-0.2491	4.32E-05
21	rs2833763	G	0.9428	0.2309	0.4902	1.395	4.50E-05
22	rs116919083	A	-1.91	0.4657	-2.822	-0.9968	4.16E-05

그림

Figure 1. Manhattan plot of lung function, KoGES

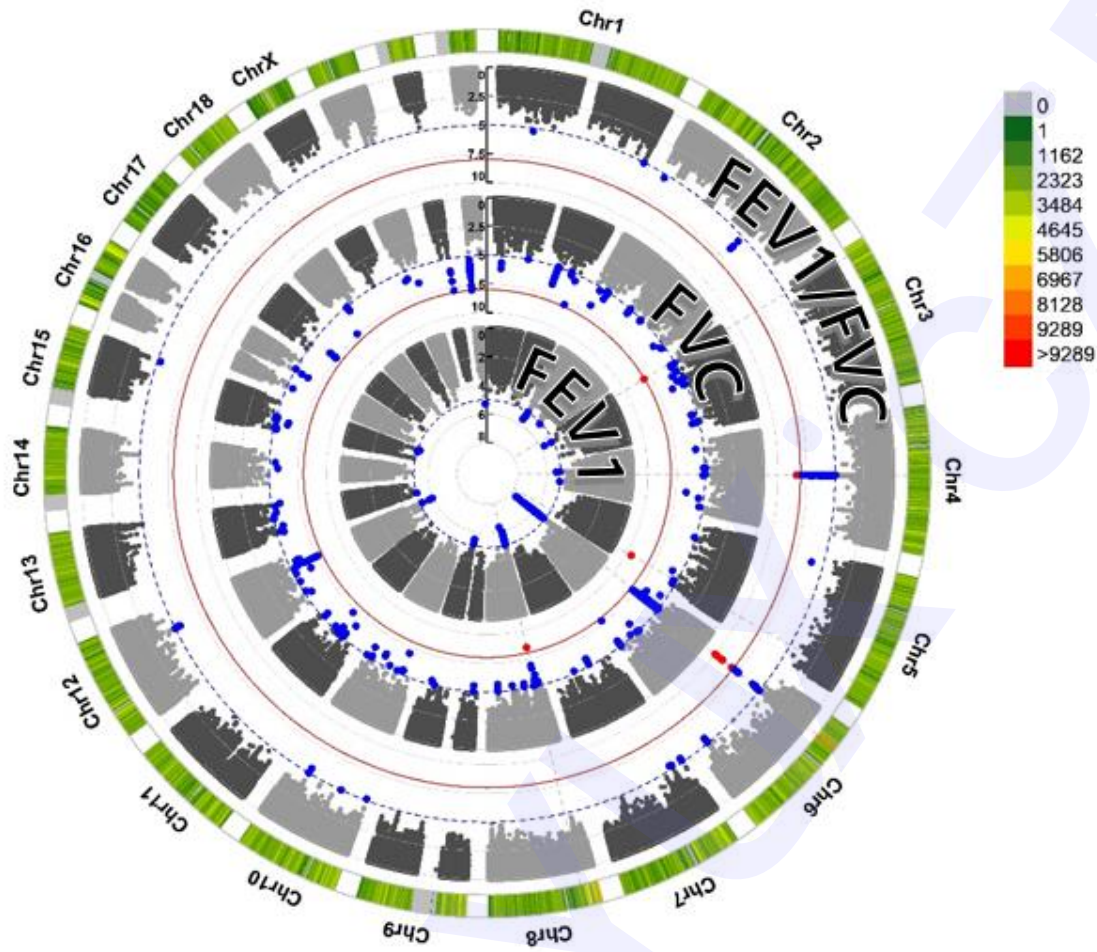


Figure 2. Manhattan plot of FEV₁/FVC ratio, KoGES

