Genetic epidemiology of pulmonary function

Yue Chen

Respiratory diseases are major threats to human health. It is believed that chronic obstructive pulmonary disease (COPD) is both environmental and genetic; however, specific genetic factors in the development of COPD have not been clearly identified, except for protease inhibitor types. Alpha1-antitrypsin deficiency is rare in the general population and accounts for less than 2% of the cases of COPD. The molecular genetics of COPD has recently been reviewed by Barnes.

Pulmonary function measures are the most important phenotypes of COPD. The respiratory muscles, the thorax, and the lungs are the components of the ventilatory apparatus which can be evaluated by appropriate function tests. The dynamic functional capacity of the ventilatory apparatus can be assessed by the volume-time or flow-volume manoeuvres. Airflow rates and volumes inhaled or exhaled over specific time intervals provide information on the flow resistive properties of the airways. These pulmonary function measures predict the development of lung diseases and overall mortality. While the environmental determinants of pulmonary function have been extensively studied—for example, smoking and ambient air pollution—the genetic determinants have recently received increasing attention. Genetic epidemiological studies of pulmonary function are of potential importance in understanding normal pulmonary function and the aetiology and prevention of COPD and other respiratory diseases. This paper reviews the familial aggregation and segregation of pulmonary function and presents evidence for different influences of heredity on airway function, lung volume, and airway-parenchymal dysanapsis (relative airway size to lung size). Some methodological issues related to segregation analysis are also discussed.

Family aggregation

FAMILY STUDIES

A number of family studies have provided evidence for familial resemblance of pulmonary function measures. Studies have consistently shown significant parent-offspring and sibling-sibling correlations in lung volume and flow rate measures (Table 1). Most studies found that spousal correlations in pulmonary function were trivial, although Higgins and Keller reported a small but significant correlation in forced expiratory volume in one second (FEV1) between spouses, and Kauffman et al found significant spousal correlations in residual forced vital capacity (FVC), FEV1, and forced expiratory flow between 25% and 75% of the vital capacity (FEF25–75%).

One study found that familial correlations in pulmonary function were dependent on familial resemblance of body habitus and were no longer significant after the ponderal index (height/weight3/2) was taken into consideration. Other studies found that the familial correlations remained significant after adjustment for both height and weight or adjustment for height alone. Over-adjustment could

### Table 1  Family studies in pulmonary function

<table>
<thead>
<tr>
<th>Reference</th>
<th>Location and year</th>
<th>Study subjects</th>
<th>Lung function indices</th>
<th>Correlations</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tager et al</td>
<td>USA, 1976</td>
<td>148 households, 469 subjects</td>
<td>FEV1, %, FEV1 score</td>
<td>0.04, 0.05 0.18*, 0.11 0.25* 0.19* 0.26*</td>
<td>FEV1 (%) less correlated than FEV1 score</td>
</tr>
<tr>
<td>Schilling et al</td>
<td>USA, 1977</td>
<td>376 families, 816 children and their parents</td>
<td>rFVC, rFEF25–75%, rVmax50</td>
<td>0.07, 0.07, 0.18* 0.15* 0.12* 0.16*</td>
<td>Separate parent-offspring correlations; smaller ones are selected</td>
</tr>
<tr>
<td>Kauffman et al</td>
<td>France, 1989</td>
<td>945 families, 1160 children and their parents</td>
<td>rFVC, rFEV1</td>
<td>0.18* 0.26 0.19 0.30*</td>
<td>Significant spousal correlations</td>
</tr>
<tr>
<td>Coultas et al</td>
<td>USA, 1991</td>
<td>733 households, 336 spouse pairs</td>
<td>rFVC, rFEF25–75%</td>
<td>0.11 0.26  0.21 0.16 0.27 0.37*</td>
<td>Parent-offspring correlations: 6–17 year group</td>
</tr>
<tr>
<td>Chen et al</td>
<td>Canada, 1996, 1997, 1998</td>
<td>1059 parent-child pairs, 412 sibling pairs</td>
<td>rFVC, rFEV1, rVmax50/FVC</td>
<td>0.10, 0.08, 0.04, 0.11 0.18 0.18 0.11* 0.15* 0.22 0.19 0.16* 0.25 0.27*</td>
<td>First degree relative correlations were not significantly different for rFEV1, rFEF25–75%, and rVmax/FVC</td>
</tr>
<tr>
<td>Givelber et al</td>
<td>USA, 1998</td>
<td>1408 families, 5003 adult subjects</td>
<td>rFEV1</td>
<td>0.05 0.19 0.12 0.22*</td>
<td>Sib-sib correlation was greater than parent-offspring correlation</td>
</tr>
</tbody>
</table>

FEV1 = forced expiratory volume in one second; FVC = forced vital capacity; FEF25–75% = forced expiratory flow between 25% and 75% of the vital capacity; Vmax50 = maximal expiratory flow rate at 50% of vital capacity; p_s = spousal correlation; p_m = mother-offspring correlation; p_f = father-offspring correlation; p_sib = sibling-sibling correlation; r = residual.

*p<0.05.
be one reason for the disappearance of familial aggregation of pulmonary function,22 and different analytical methodology could be another.

Chen et al23–25 compared the mother-offspring, father-offspring, and sibling-sibling correlations in various pulmonary function measures. No significant differences were found in the correlations of airway function measures including the FEV1, FEF25–75%, and maximal expiratory flow rate at 50% of vital capacity (Vmax50). For FVC, however, the sibling-sibling correlation was greater than the parent-sibling correlation, which was consistent with the results from other studies.24–26

This additional resemblance between siblings may be due to shared sibling environment and environmental factors may have different impacts on the lung volume measure than on the flow rate measures.28 Givelber et al26 found that the sibling-sibling correlation in FEV1 was greater than parent-offspring correlations. Since the FEV1 was measured much earlier for the parents (1948–52) than for the offspring (1971–74), measurement error could be larger for data from the parents than those from the offspring due to outmoded spirometric techniques.26 The results are not consistent in terms of differences between mother-offspring and father-offspring correlations for various pulmonary function measures. Givelber et al26 suggested a greater mother-offspring correlation in FEV1 compared with the father-offspring correlation. Coultas et al29 and Chen et al23–25 found that there was no significant difference between mother-offspring and father-offspring correlations.

### Genetic heritability

Both family and twin studies have clearly shown that pulmonary function including flow rate and lung volume measures are familial. The reason for the familial aggregation can be environmental, genetic, or both. Genetic heritability, which is the proportion of the genetic variance to the total phenotypic variance in a defined population, can be used to quantify the degree of genetic contributions. The genetic variance can be further divided into additive genetic variance, dominance variance, and epistatic variance (interlocus interaction). Heritability in the narrow sense, which is the proportion of additive genetic variance to the total phenotypic variance, is used to measure possible genetic effects not due to major gene segregation.

A number of family studies have examined the degree to which the observed familial aggregation of pulmonary function is attributable to genetic factors and have shown a moderate degree of heritability for various pulmonary function measures (table 3). Based on the data of 439 subjects from 108 families of patients without pulmonary disease, Astemborski et al30 found that additive genetic variation accounted for 28% of the variation in

![Table 2 Twin studies in pulmonary function](image-url)
residual FEV1 and 24% of the variation in residual FEV1/FVC. The estimators were smaller among families of patients with airway obstruction disease.40 Four studies used the path analysis approach to identify hereditary and environmental sources of familial aggregation for pulmonary function traits. Lewitter et al41 studied 404 nuclear families including 602 parents and 756 children and found that 42–47% of the variability in FEV1 and FEF25–75% could be explained by underlying genetic differences among the individuals. Another analysis by Cotch et al42 showed a similar estimate of heritability of 36–40% for FEV1 and no significant difference between white and black individuals. Coultas et al43 found an increased genetic variance in smokers. The genetic variances for FVC and FEV1 were 10% and 25% greater for smokers than for non-smokers, respectively.44 The heritability estimates of pulmonary function measures were lower in another study. Based on the data from 305 men and 339 women, Devor and Crawford45 estimated that approximately 20% and 17% of the variation in FVC and FEV1, respectively, was due to the transmission from parents to offspring. Smoking and age could alter the familial aggregation of pulmonary function, and different study designs may be the reason for the discrepancy.46 Shared heritable factors might not only influence the lung growth and development, but also the decline in pulmonary function in adults.47 In the recent Humboldt Family Study of 309 nuclear families, Chen et al48,49 used a class D regressive model and estimated the additive genetic heritability as 26–40% for FEV1, Vmax50, FEF25–75%, and Vmax50/FVC.

The heritability estimates for different measures of pulmonary function in twin studies have been inconsistent. Hubert et al50 studied 127 MZ and 141 DZ male twin pairs aged 42–56 years and estimated the heritability to be as high as 77% for FEV1. In a study of 256 MZ and 158 DZ adult twins, Redline et al51 found that 40–75% of the measured variability in pulmonary function was accounted for by genetic influences. In another study, however, Ghio et al52 found that the heritability was not significant after adjustment for height in 74 university student pairs of twins with an average age of 20 years. Twin studies usually provide an inflated estimate of heritability43 because twins share very similar, if not the same, exposures in utero, and share a more homogeneous living environment than other individuals. Since heritability is the ratio of the genetic variance to the total variance, heritability increases with decreasing variance due to environment. The heritability estimates can be more biased if the effects of environmental factors are more similar in MZ twins than in DZ twins. Some studies have shown that MZ twins have a greater concordance in smoking habits than DZ twins.42,43 Correlational analysis of the distribution of given traits within family members is one way to increase the robustness of the twin data analysis.44 Heritability is a population-specific parameter and is affected by the environment in which the population developed. In addition, if there is an interaction between genotype and environment—for example, smoking may alter the genetic effects on pulmonary function—it is almost impossible to separate the genetic variance and environmental variance completely. Because of these limitations, heritability estimation should be explained with caution.

### Major genetic effects on pulmonary function

#### SEGREGATION ANALYSIS

Genetic effects may be the consequence of a single gene (a major gene), a small number of genes (oligogenes), or a large number of genes each with a small effect (polygenes). In segregating families the relatively large effects of major genes should be detectable using the tools of segregation analysis, while the more general predictions of polygenes predict the overall patterns of correlation among relatives. The classical segregation analysis is used to identify Mendelian ratios when a phenotype is controlled by a major gene, which is traditionally assumed to result from segregation at a single locus having two alleles, A and B. The likelihood method is frequently used in

<table>
<thead>
<tr>
<th>Reference</th>
<th>Location and year</th>
<th>Study subjects</th>
<th>Lung function indices</th>
<th>Additive genetic heritability</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lewitter and Tager50</td>
<td>USA, 1984</td>
<td>404 families, 602 parents, 756 children</td>
<td>FEV1 score, FVC, rFEV1</td>
<td>42–47%</td>
<td>Path analysis; consistent over time</td>
</tr>
<tr>
<td>Devor and Crawford50</td>
<td>USA, 1984</td>
<td>96 families, 307 subjects</td>
<td>FEF25–75%, rFEV1, rFVC</td>
<td>20%</td>
<td>Path analysis</td>
</tr>
<tr>
<td>Astemborski et al52</td>
<td>USA, 1985</td>
<td>108 families, 439 adults</td>
<td>rFEV1, rFVC, Vmax50</td>
<td>28%</td>
<td>Variance components analysis; adult study population</td>
</tr>
<tr>
<td>Beaty et al53</td>
<td>USA, 1987</td>
<td>158 families, 781 subjects</td>
<td>rFEV1, FEF25–75%, rFEV1/FVC</td>
<td>24%</td>
<td>Variance components analysis; ascertainment through a proband with obstructive lung disease</td>
</tr>
<tr>
<td>Cotch et al54</td>
<td>USA, 1990</td>
<td>384 families, 978 subjects</td>
<td>rFEV1 (cross sectional), rFVC, rFVC (non-smoking parents)</td>
<td>36%</td>
<td>Path analysis; no inter-generational differences</td>
</tr>
<tr>
<td>Coultas et al55</td>
<td>USA, 1991</td>
<td>733 households, 336 spouse group pairs, 1099 parent-child group pairs, 412 sibling pairs</td>
<td>rFVC (non-smoking parents), rFVC (smoking parents)</td>
<td>43%</td>
<td>Path analysis; no substantial changes based on age and smoking status</td>
</tr>
<tr>
<td>Chen et al56</td>
<td>Canada 1996, 1997, 1998</td>
<td>309 families, 1045 subjects</td>
<td>rFEV1, rFVC, rFEV1 (non-smoking parents), rFEF25–75%, rVmax50/FVC</td>
<td>26%</td>
<td>Class D regressive model</td>
</tr>
</tbody>
</table>

FEV1 = forced expiratory volume in one second; FVC = forced vital capacity; FEF25–75% = forced expiratory flow between 25% and 75% of the vital capacity; Vmax50 = maximal expiratory flow rate at 50% of vital capacity; r = residual.
segregation analyses, and other approaches are also proposed including generalised estimating equations. There are various strategies of model evaluation. Figure 1 gives an example of the analytical strategy for a segregation analysis. If a pulmonary function phenotype is familial, the next step is to determine mixtures of phenotypic distribution. Before there is evidence of Mendel transmission, a more general term called “osiotypes” or “types” has been suggested to describe the mixtures of distribution. A key assumption is that individuals represent different osiotypes or essential types that may reflect different genotypes for Mendelian models. The parameters of transmission probabilities can therefore be estimated, which are the probabilities of a parent transmitting the A allele to an offspring. Under Mendelian transmission, τ(AA) = 1, τ(AB) = 0.5, and τ(BB) = 0. A non-transmitted environmental effect was obtained with the three transmission probabilities being equal (τ(AA) = τ(AB) = τ(BB)). The process of segregation analysis involves testing a series of models of inheritance, including Mendelian models such as dominant and recessive models plus non-genetic models, to identify the best fitting and most parsimonious model for a given set of family data.

AIRWAY FUNCTION

Based on data from 85 families with COPD and 56 families without pulmonary disease, Rybicki et al used the class A regressive model and found that there were major genetic effects on FEV, and the major gene effects could explain all of the familial correlations for FEV, in families ascertained through a COPD proband. In families of patients without pulmonary disease, however, there were no familial correlations for FEV, and therefore no evidence of genetic control of FEV,. These results suggested substantial aetiological heterogeneity in the control of FEV, between the families with COPD and those without the disease. The reasons for the lack of familial correlations in the families of those without COPD and its discrepancy with other studies were not discussed in the report. The class A regressive model makes an assumption that siblings are correlated only through common parentage. This restriction, in the absence of a major gene, may lead to false inference of a major gene. Chen et al studied 309 young families and used the class D regressive model in the segregation analysis, which allows additional correlation among siblings and is characterised by equal sibling-sibling correlations. The data have suggested that FEV, FEF25–75%, and Vmax50 are more likely to be controlled mainly by multiple loci—namely, many independent genes—each contributing in an additive fashion and/or common environmental factors are responsible for the familial resemblance of airway function. In a recent report of 5003 subjects from 1408 families in the Framingham study Givelber et al provided consistent results. The most parsimonious model for FEV, included non-transmitted major types and residual familial correlations, and the Mendelian hypothesis was rejected. Based on the data from 309 families (1163 individuals) in the Tuscan children’s respiratory study Holberg et al also suggested
polygenic control of FEV₁ or common environmental factors resulting in the familial aggregation of the trait; however, genetic heterogeneity might exist between families with and without asthmatic members.

Silverman et al. studied α₁-antitrypsin deficient individuals in 44 nuclear families and found that there was an additional major gene other than the Pi locus influencing FEV₁. However, the major gene effect diminished after adjustment for pack-years of smoking.

**LUNG VOLUME**

Based on the data from 309 nuclear families, Chen et al. performed a segregation analysis for FVC, a measure of lung volume. Models with both major types and familial correlations gave the best fit for the data. However, neither Mendelian nor parent-offspring transmission hypotheses were rejected. Heterogeneity may exist between families. The authors calculated the likelihood under the Mendelian model (L(Mendelian)) and the environmental model (L(environmental)) and used the ln-likelihood values to sort families into groups that support one model of inheritance over another. In a subset of 196 families with a ln(L(Mendelian)/L(environmental)) value greater than zero, the families suggested a Mendelian gene leading to lower values of FVC, and the single locus explained all the familial aggregation of residual FVC. In the other 113 families in whom the ln(L(Mendelian)/L(environmental)) value was less than zero the Mendelian hypothesis could not be rejected and the Mendelian model showed that a single locus accounted for all familial correlations except for the sibling-sibling correlation. However, the environmental hypothesis could not be rejected for this subgroup of families although the Mendelian model had a better fit than the environmental model based on the values of the Akaike’s information criterion (AIC, see later). The approach of dividing the families into two groups based on an individual likelihood ratio might remove certain confounding effects and increase the statistical power of detecting a major gene effect.

**AIRWAY-PARENCHYMAל DYSANAPSIS**

Green et al. found a low correlation between lung volume and maximal expiratory flow, and no obvious relationship between static lung recoil and Vmax₅₀, suggesting that there are substantial differences between individuals in airway size and function that are independent of lung size. Disproportionate but physiologically normal growth of airway and parenchymal components suggests a “dysanaptic” growth which may have an embryological basis. The concept of airway-parenchymal dysanapsis was advanced by Mead who reasoned that subjects with large lungs do not necessarily have larger airways than those with small lungs. He used Vmax₅₀/(VC × Pst(L)₅₀) as an index of airway-parenchymal dysanapsis in which Vmax₅₀ is the maximal expiratory flow rate at 50% of total volume and Pst(L)₅₀ is the maximal flow static recoil pressure characteristic at 50% of vital capacity (VC). Green et al. have shown that lung static recoil contributes little to the variability between individuals and that the major variability in maximum flows is attributable to airway dimensions. The correlation between Vmax₅₀/VC and Vmax₅₀/(VC × Pst(L)₅₀) is high, ranging from 0.78 to 0.84.

The dysanapsis is a general phenomenon. Airway-parenchymal dysanapsis has been observed both in adults and children. Martin et al. found substantial interindividual variability of maximal expiratory flow rates relative to lung volumes during early childhood which remained constant during growth, suggesting that the dysanapsis originates in early childhood.

Chen et al. examined the major gene effects on the ratio of Vmax₅₀ to FVC. There was evidence for mixtures of distribution while the polygenic and sporadic model did not give a good fit to these data. The transmission of osiotypes for Vmax₅₀/FVC was not different from the Mendelian expectation, and the no parent-offspring transmission hypothesis was rejected, suggesting that there is a single locus gene or a cluster of genes working in unison to determine Vmax₅₀/FVC.

It has been suggested that airway-parenchymal dysanapsis might have relevance for the pathogenesis of obstructive airway disease. Conrara et al. indicated that airway-parenchymal dysanapsis, as measured by MMFR/FVC, was a significant predictor for the degree of bronchial hyperresponsiveness.

**METHODOLOGICAL ISSUES**

**Adjustment for covariates**

Pulmonary function phenotypes are most likely to be multifactorial, controlled by both genetic and environmental factors. Various factors including host characteristics, environmental factors, and history of respiratory symptoms and disease influence these pulmonary function measures. Adjustment for these variables is always a challenge. One approach is to include these variable covariates in regressive models; however, most of the variables tend to have inconsistent effects on pulmonary function measures in different age and sex groups. For example, body weight shows both “muscularity effect” (increase in pulmonary function with increasing weight) and “obesity effect” (decrease in pulmonary function with increasing weight). There is more “muscularity effect” than “obesity effect” in children and young adults but more “obesity effect” than “muscularity effect” in older adults, which is sex related. Cigarette smoking has a sex related effect on pulmonary function measures.

Age itself positively predicts pulmonary function in children and young adults and negatively predicts pulmonary function in middle aged and older adults. Clearly, these variables cannot be appropriately adjusted by including them in the same regressive models in a segregation analysis. Preadjustment for the variables in different age and sex groups is therefore preferred. The adjusted values are used to fit models of inheritance. The relative importance of the covariates in relation to pulmonary function phenotype varies across age and sex.
groups, and this is reflected by the proportion of variation explained by these factors. Most studies of pulmonary function preadjusted these covariates including smoking, one of the most important determinants, in the segregation analyses. 23–26 48 50

It is debatable whether or not to adjust for history of respiratory symptoms and disease. A history of respiratory symptoms and disease may reduce pulmonary function but it can also be a surrogate measure for the effects of smoking and other environmental factors on the respiratory system. Adjustment for these variables may eliminate some confounding effects but may also reduce the variance of pulmonary function phenotypes unnecessarily. Comparisons of adjusted and unadjusted results are always helpful.

Selection of parsimonious models
The likelihood ratio test is usually used to select the most parsimonious model, which is minus twice the difference in the log, likelihood (ln L) between models before and after reducing parameters. The test is based on a comparison of strictly hierarchical models. For several alternative non-hierarchical models the better fitting model is considered with a lower value of the Akaike’s information criterion (AIC = −2 × ln L + 2 × number of parameters estimated). 45 Although the AIC is not a statistical test and therefore provides no statistical inference, it is useful in identifying the most parsimonious model.

Statistical power
In segregation analysis a series of models of inheritance are fitted and the most parsimonious one chosen is to explain the familial aggregation of a pulmonary function phenotype. Tests are typically based on “goodness of fit” measures and a type II error occurs when the genetic model is incorrect, but statistical testing fails to reject it because of small sample size. 29 Statistical power for segregation analysis is related to the size of gene effect and sample size. 44 Nuclear families with larger sibships are generally more informative; however, the total number of subjects rather than sibship size per se may have more influence on the power. 46 Large sample size increases the power in discriminating the completing model.

Gene-environmental interaction
A previous study has documented gene-environment interactions in COPD. 68 Another study has suggested that a gene-environment interaction may influence pulmonary function. 47 The pulmonary function phenotype expression of a gene may therefore depend on environmental variables such as smoking. Ignoring gene-environment interactions may result in underestimating the genetic effects on quantitative traits. 48

Conclusions and future directions
Both family studies and twin studies have shown familial aggregation of various measures of pulmonary function. There is a moderate degree of genetic heritability for these pulmonary function measures. However, genetic factors may have different influences on phenotypes of airway function, lung volume, and airway-parenchymal dysanapsis. Airway function phenotypes are more likely to be controlled by many loci with no major gene effects and/or are due to common environmental factors in “normal” families. Aetiological heterogeneity may exist in families with COPD or asthma and heredity may have different effects on normal airway function and airway dysfunction. There is evidence of major gene control of phenotypes of airway-parenchymal dysanapsis and lung volume. It would be interesting to investigate further the different effects of heredity on various pulmonary function phenotypes and their potential linkage to diseases of the lung. In particular, researchers should seek biological evidence for major gene controls of lung volume and airway-parenchymal dysanapsis of the lung.

This work was supported in part by the National Health Research and Development Program, Health Canada, through a National Health Research Scholar award to the author.

References


Genetic epidemiology of pulmonary function

Yue Chen

Thorax 1999 54: 818-824
doi: 10.1136/thx.54.9.818

Updated information and services can be found at:
http://thorax.bmj.com/content/54/9/818

These include:

References
This article cites 63 articles, 11 of which you can access for free at:
http://thorax.bmj.com/content/54/9/818#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Airway biology (1100)
- Lung function (773)
- Air pollution (110)
- Drugs: infectious diseases (968)
- Environmental issues (253)
- Epidemiologic studies (1829)
- Health education (1223)
- Smoking (1037)
- Tobacco use (1039)
- Genetic screening / counselling (88)
- Molecular genetics (211)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/