Genotype-phenotype correlation for pulmonary function in cystic fibrosis

J de Gracia, F Mata, A Álvarez, T Casals, S Gatner, M Vendrell, D de la Rosa, L Guarner, E Hermosilla

Background: Since the CFTR gene was cloned, more than 1000 mutations have been identified. To date, a clear relationship has not been established between genotype and the progression of lung damage. A study was undertaken of the relationship between genotype, progression of lung disease, and survival in adult patients with cystic fibrosis (CF).

Methods: A prospective cohort of adult patients with CF and two CFTR mutations followed up in an adult cystic fibrosis unit was analysed. Patients were classified according to functional effects of classes of CFTR mutations and were grouped based on the CFTR molecular position on the epithelial cell surface (I–II/I–II, I–II/III–V). Spirometric values, progression of lung disease, probability of survival, and clinical characteristics were analysed between groups.

Results: Seventy-four patients were included in the study. Patients with genotype I–II/I–II had significantly lower current spirometric values (p < 0.001), greater loss of pulmonary function (p < 0.04), a higher proportion of end-stage lung disease (p < 0.001), a higher risk of suffering from moderate to severe lung disease (odds ratio 7.12 (95% CI 1.3 to 40.5)) and a lower probability of survival than patients with genotype I–II/I–II, I–II/IV and I–II/V (p < 0.001).

Conclusions: The presence of class I or II mutations on both chromosomes is associated with worse respiratory disease and a lower probability of survival.

Cystic fibrosis (CF) is the most common recessively inherited disease in white people, occurring in approximately 1:5500 live births in our area. Patients with CF have clinical phenotypes that mainly include chronic lung infection, gastrointestinal tract alterations, and infertility in men. They have apparently normal lungs at birth, but progressive deterioration in pulmonary function is the cause of death in 95% of cases and represents the principal prognostic factor in these patients.

Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene which encodes a protein expressed in the apical membrane of exocrine epithelial cells. CFTR functions principally as a cAMP induced chloride channel and appears capable of regulating other ion channels. Mutations in the CFTR gene cause inspissated secretions leading to disease in the affected organs. Since the CFTR gene was cloned in 1989, over 1000 mutations in this gene have been identified. With reference to chloride transport dysfunction, the CFTR mutations can be grouped into five classes: (I) CFTR not synthesised, (II) defective processing, (III) defective regulation, (IV) defective conductance, (V) partially defective production or processing. This classification makes it possible to predict the likely effect of a known mutation on the CFTR function, although the effect of a given mutation on cell function may not be known in full. While a major fraction of the CFTR protein does not reach the epithelial cell surface in the presence of mutation classes I and II, it is present on the cellular surface in mutation classes III, IV or V, and a certain residual function can be found. This variation in the genotype provides a rationale for effects of the CFTR mutations on phenotype.

The relationship between genotype and congenital bilateral absence of the vas deferens and the relationship between genotype and pancreatic insufficiency have been established in several publications. Although there are a few rare mutations such as A455E and R117H which are clearly linked to a better pulmonary outcome, the effect on the lungs of F508del and most other mutations cannot be separated and attempts to link mutations in CFTR to severity of lung disease have been unsuccessful. Furthermore, genes other than CFTR and environmental factors such as access to specialised centres and treatment strategies may be more important factors in modifying the development, progression, and severity of pulmonary disease. Possible reasons for the lack of correlation between genotype and pulmonary disease include: (1) the majority of studies have been carried out in children and young patients; (2) the relatively short evolution of the disease in these patients; (3) the more effective treatment against rapid progression of lung damage in the first years of life; and (4) a lack of studies which include mutations found most frequently when a diagnosis of CF is established in adulthood.

The hypothesis of this study was that the evolution of pulmonary disease and the probability of survival may be related to whether or not the CFTR protein reaches the epithelial cell surface and a certain residual function could be present. A prospective study of a cohort of adult patients treated and followed up at the same CF unit was performed.

METHODS
Adult patients (>16 years) diagnosed as having CF with known genotype included in CF Mutation Database (Genetic Analysis Consortium) and followed up in the Adult CF Unit of our centre between January 1992 and December 2002 were recruited.

Study design
A prospective cohort study was performed to investigate the relationship between genotype and progression of lung disease. The primary end points were decrease in pulmonary function.
function and mortality from pulmonary disease. The trial was approved by the hospital ethics committee.

A clinical evaluation, sputum cultures, and pulmonary function test were carried out at each medical check-up every 3 months and whenever necessary during follow up. Blood and urine tests were analysed every 6 months. Thoracic and abdominal CT scans were performed at the time of diagnosis and biennially. Chronic bronchial colonisation was considered when three or more sputum cultures were persistently positive over a period of 6 months. Pancreatic insufficiency was assessed by fecal fat and/or fecal elastase levels in all patients. Additional information was obtained by CT scans of the pancreas and nutritional status was determined according to sex, age, and body mass index (weight (kg)/height$^2$ (m$^2$)). The demographic and clinical characteristics of the patients analysed were those available on their last visit to the unit.

### Pulmonary function tests

Forced vital capacity (FVC % predicted) and forced expiratory volume in 1 second (FEV1 % predicted) were considered only if patients were clinically stable (absence of pulmonary exacerbation over the previous 4 weeks). Predicted values for forced spirometry were taken from Roca et al.; 60% of the predictive value was taken as a cut-off value to differentiate between a moderate abnormality and a moderate to severe abnormality. The pulmonary function test results obtained on the first visit to the adult unit were considered as the baseline values and those obtained on their last visit to the unit as the current values. A decrease in pulmonary function was calculated and adjusted for age and time of follow up.

### Genetic study

Molecular analysis of the CFTR gene included the detection of the 31 most common mutations (Genotype Cystic Fibrosis Diagnosis System; PE Applied Biosystems, CA, USA). A wider genetic study was carried out if necessary in the molecular genetics department of the Oncology Research Institute, Duran y Reynals Hospital, Barcelona as previously described. The whole coding region and intronic boundaries of the CFTR gene were analysed using multiplex denaturing gradient gel electrophoresis (DGGE) and single strand conformation polymorphism analysis (SSCP/Heteroduplex; Genephor, Amersham Pharmacia Biotech, Buckinghamshire, UK). The combination of these techniques gives a mutation detection level of 97% in the Spanish CF population (T Casals, unpublished data). The fragments with an abnormal migration pattern were characterised by sequencing using the BigDye Terminator Cycle Sequencing kit (PE Applied Biosystems; PE Applied Biosystems, CA, USA). A wider range of mutations was detected.

### Table 1: CFTR mutation according to functional classification

<table>
<thead>
<tr>
<th>Class</th>
<th>Molecular dysfunction</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Defective protein processing</td>
<td>F508del, N1303K, IS07del, R1066C</td>
</tr>
<tr>
<td>III</td>
<td>Defective protein regulation</td>
<td>D1270N, 5G→A, 1811+1.6kbA→G, 3849-10kbC→T, 3272+26G→A</td>
</tr>
<tr>
<td>V</td>
<td>Partially defective production</td>
<td>2789+5G→A, 1811+1.6kbA→G, 3849-10kbC→T, 3272+26G→A</td>
</tr>
</tbody>
</table>

### Table 2: Groups based on genotype in CF adult patients

<table>
<thead>
<tr>
<th>Functional classes</th>
<th>Genotype</th>
<th>No of subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>I–I</td>
<td>G542X/W1282X, R1162X/1898+3A→G, R1162X/CFTRdelen19</td>
<td>1</td>
</tr>
<tr>
<td>I–III</td>
<td>G542X/D1270N+R74W, 711+1G→T, G551D</td>
<td>1</td>
</tr>
<tr>
<td>I–IV</td>
<td>G542X/P205S, Q890X/1609delCA, 1.6kbA</td>
<td>1</td>
</tr>
<tr>
<td>I–V</td>
<td>G542X/2789+5G→T, G542X/1811+1.6kbA→G, 1782delA/2789+5G→A, 1609delCA/1811+1.6kbA→G</td>
<td>1</td>
</tr>
<tr>
<td>II–III</td>
<td>F508del/F508del, N1303K/1609delCA, R1066C</td>
<td>21</td>
</tr>
<tr>
<td>II–IV</td>
<td>F508del/D1270N+R74W, IS07del/D1270N+R74W, 609delCA</td>
<td>1</td>
</tr>
</tbody>
</table>

### Statistical analysis

Descriptive statistics were calculated for continuous variables and frequency statistics for categorical variables. Exploratory data analysis (EDA): histogram, box plot, density plot and normal probability-probability plots (pp-plot) were used for visual normality examination of the continuous variables. Differences between means in categorical variables were performed with the ANOVA method. To study the decline in pulmonary function between groups the ANOVA method (repeated measures) was used with baseline and current spirometric values as dependent variables, genotype groups as the independent variable, and age and evolution time as the independent variable, and age and evolution time as the independent variable.
adjusted variables. ANOVA method (two way) and Scheffe test as post hoc test were carried out to evaluate the association between classes of mutation on both chromosomes (I–II/I–II, I–II/III, I–II/IV, I–II/V and I–II/III–V) and pulmonary function tests. Fisher’s exact test was used to study differences between proportions. Survival was analysed by Kaplan-Meier survival plots and log rank tests; the time to an event was calculated from the date of birth to the date of end stage lung disease (date of lung transplantation or date of death secondary to lung disease). Univariate and multivariate logistic regression models were used to examine the relationship between pulmonary function and both genotype and pancreatic insufficiency. The significance level was p<0.05 (two tailed). All tests were performed using SPSS software Version 11.0 (SPSS, Chicago, IL, USA).

RESULTS

A total of 120 adult patients diagnosed as having CF were followed up in our CF unit during the study period. Seventy four (40 men) of mean (SD) age 26.7 (8.3) years in whom two CFTTR mutations could be detected were included in the study. Forty four patients (59.5%) were transferred from the paediatric CF unit and the other 30 (40.5%) were diagnosed in our adult CF unit. The mean (SD) time of follow up in the paediatric CF unit and the other 30 (40.5%) were diagnosed in our adult CF unit during the study period. Seventy patients (58%) were diagnosed as class I or II mutations on both chromosomes (log rank test for trend p<0.001, fig 3). The correlation study revealed a significant relationship between the pair of mutations and severity of pulmonary disease. Patients with genotype I–II/I–II had a higher risk of developing moderate to severe pulmonary disease adjusted for age at diagnosis and time of follow up (p<0.04, fig 2).

A survival study was carried out and the probability of suffering from end-stage lung disease was significantly higher among patients with class I or II mutations on both chromosomes (log rank test for trend p<0.001, fig 3). The correlation study revealed a significant relationship between the pair of mutations and severity of pulmonary disease. Patients with genotype I–II/I–II had a higher risk of developing moderate to severe pulmonary disease adjusted for pancreatic insufficiency (odds ratio (OR) 7.1 (95% CI 1.3 to 40.5) in univariate and multivariate analysis, table 3). The demographic and anthropometric characteristics as well as the pancreatic and respiratory function status of the groups are shown in table 4. Patients with class I or II

The groups based on the CFTTR mutation functional classes on both chromosomes are shown in table 2.

There were no significant differences in the mean (95% CI) current FVC and FEV1 predicted values between patients with genotypes I–II/I–II, I–II/IV and I–II/V, but these values were significantly higher than those observed in patients with genotype I–II/I–II (fig 1).

The evolution of lung disease was significantly different between patients with genotype I–II/I–II and those who had class III, IV or V CFTTR mutations on at least one chromosome. Patients with mutation class I or II on both chromosomes had lower mean baseline FVC and FEV1 predicted values and a more significant decrease in pulmonary function during follow up than patients with at least one class III, IV or V mutation. These differences persisted when progression of lung disease was adjusted for age at diagnosis and time of follow up (p<0.04, fig 2).

A survival study was carried out and the probability of suffering from end-stage lung disease was significantly higher among patients with class I or II mutations on both chromosomes (log rank test for trend p<0.001, fig 3). The correlation study revealed a significant relationship between the pair of mutations and severity of pulmonary disease. Patients with genotype I–II/I–II had a higher risk of developing moderate to severe pulmonary disease adjusted for pancreatic insufficiency (odds ratio (OR) 7.1 (95% CI 1.3 to 40.5) in univariate and multivariate analysis, table 3).

The demographic and anthropometric characteristics as well as the pancreatic and respiratory function status of the groups are shown in table 4. Patients with class I or II

The progression of lung damage was significantly higher in patients with CFTTR class I or II mutations on both chromosomes (p<0.04). Data are shown as mean (95% CI) predicted values. FVC, forced vital capacity; FEV1, forced expiratory volume in 1 second.

Figure 2. Evolution of pulmonary disease during follow up in the adult cystic fibrosis unit adjusted for age and time of evolution estimated using ANOVA. The progression of lung damage was significantly higher in patients with CFTTR class I or II mutations on both chromosomes (p<0.04). Data are shown as mean (95% CI) predicted values. FVC, forced vital capacity; FEV1, forced expiratory volume in 1 second.
mutations on both chromosomes had a significantly higher prevalence of pancreatic insufficiency, chronic bronchial colonisation with *Pseudomonas aeruginosa* or *Staphylococcus aureus* and end-stage lung disease; (p<0.001). Patients with at least one class III, IV or V mutation were older at the time of diagnosis and had higher anthropometric rates.

**DISCUSSION**

This study shows a relationship between the CFTR mutation functional class on both chromosomes and pulmonary function in adult patients with CF. The patients who had CFTR mutation classes I or II on both chromosomes showed significantly lower baseline and current spirometric values, greater loss of pulmonary function during follow up, higher risk of developing moderate to severe pulmonary disease, and a lower rate of survival due to end-stage lung disease than patients with at least one CFTR functional class III, IV or V mutation. It has previously been suggested that environmental factors constitute the most important factors in modifying the development, progression, and severity of pulmonary disease in CF. Access to specialised centres, treatment strategies, and socioeconomic status have been shown to affect the long term outcome over the last few years. In this study the impact of these factors was reasonably reduced as all patients had been treated and followed up in the same CF unit since their diagnosis was established. The National Health Service (NHS) in our country guarantees access for the whole population regardless of their socioeconomic status, all the CF units are integrated in the NHS, and treatment is provided free of charge to all patients without any restriction. Given that CF patients are born with apparently normal lungs and that lung damage progresses over time, we were able to study a relationship between genotype and pulmonary phenotype from birth to their admission to the adult unit and afterwards during follow up. Paradoxically, the baseline spirometric values on admission to the adult unit were significantly lower in patients with class I or II mutations—most of whom had received specialised treatment from the time of diagnosis in childhood—than in those with at least one class III, IV or V mutation, most of whom had received specialised treatment only after their diagnosis in adulthood. These findings suggest that genotype is more important than environment as a prognostic factor of pulmonary phenotype in CF patients. Nevertheless, the environmental features together with possible genetic modifiers could account, at least in part, for the variability of pulmonary phenotype observed in some cases in patients with the same genotype.

Pulmonary function can be maintained unimpaired or slightly impaired during the first years of life, and long periods of follow up are required to observe differences in patients with different genotypes. Most studies designed to demonstrate the genotype-phenotype correlation used populations of patients diagnosed with CF in childhood or youth. In these reports F508del and most other mutations cannot be separated with respect to their effect on the lungs, and attempts to link mutations in CFTR to severity of lung disease have not been successful. The high prevalence of class I and II mutations in patients diagnosed at an early age and the relatively short period of follow up could be critical in preventing the establishment of a relationship between genotype and pulmonary disease in those trials. Unlike other reports, all the patients in this study were adults with a higher average age and were evaluated at the same centre. 40% of these patients were diagnosed as adults and most of them had at least one class III, IV or V CFTR mutation (table 4); genotype-phenotype correlation for pulmonary function was observed. The mild pulmonary phenotype seen in patients with genotype I–II/III–V is consistent with previous reports where a genotype-phenotype correlation for pulmonary function was demonstrated in patients diagnosed with CF in childhood or youth. In these reports F508del and most other mutations cannot be separated with respect to their effect on the lungs, and attempts to link mutations in CFTR to severity of lung disease have not been successful. The high prevalence of class I and II mutations in patients diagnosed at an early age and the relatively short period of follow up could be critical in preventing the establishment of a relationship between genotype and pulmonary disease in those trials. Unlike other reports, all the patients in this study were adults with a higher average age and were evaluated at the same centre. 40% of these patients were diagnosed as adults and most of them had at least one class III, IV or V CFTR mutation (table 4); genotype-phenotype correlation for pulmonary function was observed. The mild pulmonary phenotype seen in patients with genotype I–II/III–V is consistent with previous reports where a genotype-phenotype correlation for pulmonary function was demonstrated in patients diagnosed with CF in childhood or youth. In these reports F508del and most other mutations cannot be separated with respect to their effect on the lungs, and attempts to link mutations in CFTR to severity of lung disease have not been successful. The high prevalence of class I and II mutations in patients diagnosed at an early age and the relatively short period of follow up could be critical in preventing the establishment of a relationship between genotype and pulmonary disease in those trials. Unlike other reports, all the patients in this study were adults with a higher average age and were evaluated at the same centre. 40% of these patients were diagnosed as adults and most of them had at least one class III, IV or V CFTR mutation (table 4); genotype-phenotype correlation for pulmonary function was observed. The mild pulmonary phenotype seen in patients with genotype I–II/III–V is consistent with previous reports where a genotype-phenotype correlation for pulmonary function was demonstrated in patients diagnosed with CF in childhood or youth. In these reports F508del and most other mutations cannot be separated with respect to their effect on the lungs, and attempts to link mutations in CFTR to severity of lung disease have not been successful.

The findings observed in this study support the hypothesis that differences in CF pulmonary phenotype could be related to the effect of the genotype on CFTR protein production and function. Nevertheless, it is important to recognise that specific mutations may have characteristics of more than one

### Table 3: Risk of suffering from moderate to severe lung disease in patients with class I or II mutations on both chromosomes

<table>
<thead>
<tr>
<th>Mutations on both chromosomes</th>
<th>Odds ratio* (95% CI)</th>
<th>Odds ratio† (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVC ≥60%</td>
<td>49</td>
<td>16 (33%)</td>
</tr>
<tr>
<td>FVC &lt;60%</td>
<td>25</td>
<td>21 (84%)</td>
</tr>
<tr>
<td>FEV1 ≥60%</td>
<td>37</td>
<td>8 (25%)</td>
</tr>
<tr>
<td>FEV1 &lt;60%</td>
<td>37</td>
<td>29 (78%)</td>
</tr>
</tbody>
</table>

FVC, forced vital capacity; FEV1, forced expiratory volume in 1 second.

*Calculated by univariate logistic regression models (95% CI).
†Adjusted for pancreatic insufficiency.
Previous reports have shown a relationship between genotype and the association between pancreatic insufficiency and severity of lung disease.\textsuperscript{3,4} Hence, the presence of pancreatic insufficiency was considered the most important prognostic factor of pulmonary function. However, in this report, univariate logistic regression analysis showed a significant correlation between genotype and severity of pulmonary damage which persisted when the statistical analysis was adjusted for the presence of pancreatic insufficiency. This suggests that pulmonary function is a phenotypic expression which independently predicts the prognosis of the disease.

The effective treatment of pancreatic, pulmonary, and digestive disorders has dramatically improved the survival rates of patients with CF over the last 30 years. Currently, most deaths occur in adulthood and progressive pulmonary impairment is the main cause. In this study all deaths were due to pulmonary disease and all but one patient had a class I or II mutation on both chromosomes. In these patients the probability of survival—when the time to an event was calculated from the date of birth to the development of end-stage lung disease—was lower than in those patients whose genotype included at least one class III, IV or V mutation. These results are consistent with those observed by McKone et al\textsuperscript{3} in a retrospective study using the Cystic Fibrosis Foundation National Registry. They found that patients who were homozygous for F508del have significantly higher overall mortality and higher crude mortality adjusted for sex and age than those who were homozygous for mutation class IV and V or heterozygous for F508del with R117H, S1248del and class III mutations in our study makes it difficult to reach conclusions about its correlation with pulmonary phenotype. Only the mutations D1270N and G551D were analysed and they were associated with a more favourable outcome of pulmonary function. Nevertheless, previous studies have pointed out that, among functional class III mutations, there may exist a wide variability in their phenotypes that depends principally on the CFTR protein site for which they code.\textsuperscript{5,6}

In summary, the results of this study suggest that the genotype, based on functional class mutation on the two chromosomes, seems to be one of the most decisive factors for pulmonary phenotype and for survival in relation to pulmonary damage. Patients with genotypes that include class I or II mutations on both chromosomes have more rapid deterioration in lung function and lower survival rates related with lung disease than the other genotypes, especially in those with at least one class IV or V mutation.

**ACKNOWLEDGEMENTS**

The authors acknowledge the help of Francis McCabe in the writing of the manuscript.

**Authors’ affiliations**

J de Gracia, F Mata, A Alvarez, D de la Rosa, Department of Pneumology, Hospital General Vall d’Hebron, Barcelona, Spain

F Mata, Department of Medicine, Universidad Autonoma de Barcelona, Spain

T Casals, Medical and Molecular Genetics Center, Institut Recerca Oncològica (IRO), Hospital Duran i Reynals, Barcelona, Spain

M Vendrell, Department of Pneumology, Hospital Josep Trueta, Girona, Spain

L Guarner, Department of Gastroenterology, Hospital General Vall d’Hebron, Barcelona, Spain

S Gatner, Department of Pediatrics, Hospital General Vall d’Hebron, Barcelona, Spain

E Hermosilla, Department of Biostatistics, Hospital General Vall d’Hebron, Barcelona, Spain

**REFERENCES**


---

**Table 4** Demographic and clinical characteristics of groups according to genotype\textsuperscript{*}

<table>
<thead>
<tr>
<th>Genotype</th>
<th>n = 37</th>
<th>Genotype</th>
<th>n = 37</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>I–II/I–II</td>
<td>22 (59.5%)</td>
<td>I–II/I–III, IV or V</td>
<td>22 (59.5%)</td>
<td>0.484</td>
</tr>
<tr>
<td>Mean (SD) age (years)</td>
<td>3.7 (3.7)</td>
<td>3.7 (3.7)</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>Mean (SD) BMI (kg/m\textsuperscript{2})</td>
<td>18.4 (2.9)</td>
<td>18.4 (2.9)</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Mean (SD) sweat chloride concentration (mEq/l)</td>
<td>108 (23)</td>
<td>108 (23)</td>
<td>0.001</td>
<td></td>
</tr>
</tbody>
</table>

**BMI**: body mass index.

\textsuperscript{*}Data obtained at the first visit to adult unit.
Genotype-phenotype correlation for pulmonary function in CF


Cystic Fibrosis Mutation Database. The Cystic Fibrosis Genetic Analysis Consortium. www.genet.sickkids.on.ca/cftr/.


