

Quantification of the Relative Contribution of Environmental and Genetic Factors to Variation in Cystic Fibrosis Lung Function

J. Michael Collaco, MD, Scott M. Blackman, MD, PhD, John McGready, PhD, Kathleen M. Naughton, CRNP, and Garry R. Cutting, MD

Objective To assess the relative contributions of environmental and genetic factors to variation in cystic fibrosis (CF) lung disease.

Study design Genetic and environmental contributions were quantified by use of intrapair correlations and differences in CF-specific forced expiratory volume in 1 second measures from 134 monozygous twins and 272 dizygous twins and siblings while in different living environments (ie, living with parents vs living alone), as well as by use of intraindividual differences in pulmonary function from a separate group of 80 siblings.

Results Pulmonary function among monozygous twins was more similar than among dizygous twin and sibling pairs, regardless of living environment, affirming the role of genetic modifiers in CF pulmonary function. Regression modeling revealed that genetic factors account for 50% of pulmonary function variation, unique environmental or stochastic factors (36%), and shared environmental factors (14%; $P < .0001$). The intraindividual analysis produced similar estimates for the contributions of the unique and shared environment. The shared environment effects appeared primarily because of living with a sibling with CF ($P = .003$), rather than factors within the parental household ($P = .310$).

Conclusions Genetic and environmental factors contribute equally to pulmonary function variation in CF. Environmental effects are dominated by unique and stochastic effects rather than common exposures. (*J Pediatr* 2010; ■: ■-■).

Cystic fibrosis (CF) is a monogenic disease affecting more than 30 000 individuals in the United States and more than 70 000 worldwide. The gene responsible for CF, the cystic fibrosis transmembrane conductance regulator gene (*CFTR*), was identified more than 20 years ago.^{1,2} *CFTR* genotype-phenotype studies have demonstrated that lung disease severity can vary drastically in individuals with identical *CFTR* genotypes.³⁻⁵ These observations indicate that factors besides *CFTR*, such as environmental factors and modifier genes must contribute to pulmonary function decline, the major cause of death in CF. Indeed, much of the progressive increase in the median predicted age of survival of patients with CF in the United States to 36.9 years as of 2006 can be attributed to the substantial effect of environmental modification.⁶ On the other hand, recent twin-based studies have demonstrated that genetic factors also play a role in lung disease variation.^{7,8}

Quantifying the environmental contribution to pulmonary function is important for several reasons. First, even though a number of environmental factors have been demonstrated to affect CF lung disease, including second-hand smoke exposure,⁹⁻¹³ socioeconomic status,¹⁴ health care access,¹⁵⁻¹⁷ and air pollution,¹⁸ estimates of the contribution of environmental factors to lung disease as a whole have not been provided by previous studies.^{7,8} Second, parsing the contribution of shared versus unique environmental exposures can help assess risks when patients with CF come into contact with others in settings such as clinics and camps. Third, quantifying the contributions of environmental factors relative to genetic factors in lung disease variation can inform efforts to identify gene modifiers by use of genome-wide approaches. Both genetic and environmental factors have been quantified for other Mendelian disorders, such as the age of onset of Huntington's disease.¹⁹

To estimate the relative contribution of genetic factors, we examined intrapair correlations and confirmed our findings using an intrapair difference regression. To estimate the relative contribution of environmental factors, we used the previous intrapair difference regression and then validated our findings using intraindividual difference regression analysis in a different subset of the study population. These are the first quantitative estimates of the relative contributions of environmental and genetic factors to CF lung disease variation.

ANOVA	Analysis of variance
CF	Cystic fibrosis
<i>CFTR</i>	<i>Cystic fibrosis transmembrane conductance regulator</i> gene
FEV ₁	Forced expiratory volume in 1 second

From the Eudowood Division of Pediatric Respiratory Sciences (J.C.), the Division of Pediatric Endocrinology (S.B.), the Department of Biostatistics (J.M.), and the McKusick-Nathans Institute of Genetic Medicine ((K.N., G.C.), Johns Hopkins University, Baltimore, MD

Supported by NHLBI grant HL68927 and CF Foundation funding (CUTTIN06P0 and COLLAC09A0). The authors declare no conflicts of interest.

0022-3476/\$ - see front matter. Copyright © 2010 Mosby Inc. All rights reserved. 10.1016/j.jpeds.2010.05.018

Methods

A total of 1528 individuals in 752 families were recruited through the Cystic Fibrosis Twin-Sibling Study, including 75 sets of monozygous twins, 24 sets of dizygous twins, and one set of dizygous triplets (Figure 1; available at www.jpeds.com). Subjects attended U.S. CF centers, excepting 12 families recruited from Australia and 6 from Scotland. Informed consent was obtained from all subjects or guardians under the auspices of the Johns Hopkins University IRB. Zygosity was verified with the AmpFISTR Profiler kit (Applied Biosystems, Foster City, California).

Subjects met diagnostic criteria for CF.²⁰ One hundred forty-four individuals from families with more than 2 affected siblings were excluded because of the complexity of interactions among family members. Twenty-six subjects were excluded because their sibling was not enrolled at the time of analysis. One hundred fourteen individuals were excluded because both family members of a pair lacked pulmonary function data; 152 individuals were excluded because 1 family member of these pairs lacked pulmonary function data. Of these 190 individuals lacking data, 102 were younger than 6 years old. Ten siblings were excluded because of discordant CFTR genotypes among affected family members. Two siblings were excluded because of lack of CFTR genotype data. Thus 1080 subjects in 540 families comprise the overall population from which to select family pairs for intrapair and intraindividual analyses (Table I; available at www.jpeds.com).

Intrapair analyses were conducted with all available monozygous twin pairs (n = 67 pairs) (Table II). The relative paucity of pairs of dizygous twins both affected by CF necessitated creating sibling pair proxies, similar to previous CF twin-based studies.^{7,8} Within families in the combined dizygous twins/siblings (Dizygous/Sib), siblings and twins are sex concordant (ie, both males or both females) and born within 3 years of each other to minimize cohort variation. The Dizygous/Sib group included 11 pairs of dizygous twins and 125 pairs of siblings; the mean age difference between siblings in this group was 2.1 ± 0.6 years. Intraindividual analyses for replication were conducted with 40 pairs of siblings not used in the intrapair analyses; these subjects had pulmonary function data before and after leaving the parental home.

The calendar year of leaving the parental home was ascertained via questionnaires or clinical record review for 328 of 1080 subjects; for these subjects, the mean age at leaving the parental home was 19.6 ± 2.9 years. Six hundred thirty-six subjects were younger than 18 years old at the time of their most recent pulmonary function test and were assumed to be living in the parental home. For the remaining 116 subjects, the year was assigned on the basis of when they turned 18. Living together was defined as the time period before or on the year of moving out for the elder family member; living apart was defined as the time period after the year of moving out for each individual. One thousand twenty-nine subjects had data while living together, and 365 while living apart.

Pulmonary function data were supplemented with the Cystic Fibrosis Foundation Data Registry. Raw forced expiratory volume in 1 second (FEV₁ [liters]) was converted to CF-specific percentiles (range 0%-100%) for subjects between 6 and 40 years of age; this age- and sex-corrected phenotype allows for easy comparison of subjects with CF²¹ and remains relatively stable over time (mean rate of decline is $0.0 \pm 3.0\%$ per year).⁷ For the study population, the mean standard deviation for the average CF-specific FEV₁ percentile (averaged for all pulmonary function tests available for an individual) was 13.6% (n = 1056). Tests performed after lung transplantation were excluded. Mean pulmonary function while living together, while living apart, and during discrete time periods was derived by averaging the highest percentile per calendar quarter.⁷

Statistical Analysis

Statistical methods used include Student's *t* tests, χ^2 tests, analysis of variance (ANOVA) tests with Bonferroni corrections, and linear regression models. All regression models were subject to clustering by family. The dependent variable was either an intrapair or intraindividual (absolute) difference of the CF-specific FEV₁; the specific coding of independent variables for each model is provided in the Results section. Heritability estimates were calculated as twice the difference between the monozygous twin *r* and the Dizygous/Sib group *r*.²² For these estimates, mean Pearson correlation coefficients (*r*) were determined by permuting the assignment of twins or siblings as "A" or "B" 10⁶ times to avoid bias that might occur if one twin was consistently

Table II. Characteristics of study subjects

	Intrapair analyses		Intraindividual analyses
	Monozygous twins	Combined dizygous twins/siblings	Siblings
All individuals	134 (67 pairs)	272 (136 pairs)*	80 (40 pairs)
Mean age \pm SD at most recent PFT (range)	18.9 ± 8.9 yrs (6.1-39.8 yrs)	17.4 ± 7.5 yrs (6.1-40.0 yrs)	26.0 ± 6.0 yrs (18.7-40.0 yrs)
Sex: % Male	53.7	58.1	52.5
CFTR Genotype: % Δ F508 Homozygote	62.7	49.3	57.5
Mean FEV ₁	59.3 ± 23.9	61.4 ± 23.3	63.3 ± 21.3
CF% while living together	Median: 62.2 (n = 130)	Median: 65.2 (n = 260)	Median: 65.8 (n = 80)
Mean FEV ₁	54.5 ± 30.1	63.5 ± 28.0	71.2 ± 21.1
CF% while living apart	Median: 52.6 (n = 65)	Median: 71.3 (n = 90)	Median: 77.3 (n = 80)

*Same-sex dizygous twins (n = 11 pairs) and same-sex siblings born within 3 years of each other (n = 125 pairs).

assigned as “A.” Intercooled Stata 10 (StataCorp LP, College Station, Texas) was used for all statistical analyses.

Results

Subjects in the intraindividual analyses were older on average than subjects in the intrapair analyses (t test P value $< .0001$); however, because CF-specific FEV₁ percentiles correct for age, this difference was irrelevant in our analysis. Although more of the monozygous twins were $\Delta F508$ *CFTR* homozygotes than other subjects, this was not statistically significant (χ^2 P value: $.070$). Mean pulmonary function was not statistically different among all groups while living together (ANOVA, P value = $.453$), but monozygous twins had decreased pulmonary function compared with the siblings in the intraindividual analyses (ANOVA Bonferroni P value = $.001$).

Correlating pulmonary function from 65 pairs of monozygous twins and 130 pairs of Dizygous/Sib pairs from the intrapair population yielded a heritability estimate of 0.57 ; the data for this estimate were obtained when the pairs were living together. This estimate (>0.5) suggests a substantial genetic contribution to pulmonary function variation. However, genetic effects may be overestimated because of the shared environment, and environmental effects may be underestimated because of survival biases (ie, underrepresentation of highly discordant pairs where only one individual survives). To address these issues, we examined the correlations in subjects while living together and while living apart. The monozygous twin and Dizygous/Sib groups had similar amounts of time living together in the parental home (11.4 vs 11.3 years), and living apart after leaving the parental home (5.6 vs 4.4 years). The correlations in monozygous twins ($n = 29$ pairs) versus dizygous twins and siblings ($n = 33$ pairs) before and after leaving the parental household provide information on the contribution of genetic factors versus shared environment. Pulmonary function remained highly correlated among monozygous twins after they separate into less similar environments (living together: 0.91 vs living apart: 0.87) (Figure 2; available at www.jpeds.com); dizygous twins and siblings also showed minimal change with changes in living environment (0.51 vs 0.52).

Intrapair differences in pulmonary function between pairs of monozygous twin and pairs of dizygous twins and siblings were used as an alternative means of estimating of genetic and nongenetic contributions to disease variation. The mean (absolute) difference in CF-specific FEV₁ between twin A and twin B (or sibling A and sibling B) for dizygous twins and siblings was greater than monozygous twins, regardless of living environment (Table III: together, P value = $.0006$; apart, $.0035$), but the mean differences of monozygous or Dizygous/Sib pairs living together versus living apart were not statistically different (together, P value = $.1412$; apart, $.3091$). Of note, this difference appeared to begin to increase for both groups 3 to 4 years after the first family member left the parental home; the difference in pulmonary function between monozygous twins increased by 18 percentile points per 10 years, and for dizygous twins and

siblings by 25 percentile points per 10 years, but neither of these rates is statistically different from baseline (P value: monozygous, $.075$; Dizygous/Sib, $.302$) (Figure 3; available at www.jpeds.com). Together, these analyses indicate that genetic factors continue to influence pulmonary function in twins and siblings with CF, even after living apart for 5 years.

One means to determine whether accumulated environmental exposures contribute to variation is to examine correlations between twin and sibling pairs over time. To do this, we derived correlations and heritability estimates from all eligible subject pairs from the intrapair population within a specified age range. Although heritability estimates were >0.5 for each age stratum between ages 10 years old and 30 years old, correlation (r values) between family members progressively decreased by age group for both monozygous twins and dizygous twins/siblings after age 18 (Table IV; available at www.jpeds.com). These data indicate a role for nongenetic factors, which could be due to changes in the shared environment, unique environment, or stochastic factors.

Differences in monozygous twins (sharing 100% of their genes) living together (sharing a common environment) can be attributed to effects of unique environment or stochastic factors. In this study, monozygous twins living together showed appreciable differences in pulmonary function measures (Table III; mean difference: 8.0). *CFTR* genotype was not a significant source of variation, because analysis limited to $\Delta F508$ homozygotes (Table III) obtained similar P values for the genetic contribution and a similar mean difference for monozygous twins living together (6.9). We then performed regression modeling of intrapair differences to determine the contribution of unique environmental factors and to evaluate the relative contributions of genes and environment: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ where: $Y =$ Intrapair CF-specific FEV₁ (absolute) difference; $X_1 = [0$ if living together, 1 if living apart]; $X_2 = [0$ if monozygous twins, 1 if dizygous twins/siblings]

The intercept value (β_0) was assumed to be due to unique environmental and stochastic factors, after accounting for shared environmental factors (β_1) and genetic (β_2). With the 62 pairs in Table III, it was found that $\beta_0 = 8.0 \pm 1.5$ (P value: $< .001$), $\beta_1 = 3.2 \pm 1.9$ (P value = $.100$), and $\beta_2 = 11.2 \pm 2.7$ (P value: $< .001$), with an overall model P value of $.0001$. Thus unique environmental factors accounted for 36% ($8.0/22.4$) of variation, common or shared environmental factors 14% ($3.2/22.4$) and genetic factors 50% ($11.2/22.4$). The genetic and environmental contributions appear to be of similar magnitude (3.2 [shared] + 8.0 [unique] = 11.2 [total environment] vs 11.2 [genetic]).

To verify the above estimates, we examined pulmonary function in the same individual before and after leaving the parental household (Figure 4). By comparing pulmonary function over different time frames in the same individual, genetic effects are assumed to be nil; variation in that individual is assumed to be a function of shared and unique environmental changes. This method also allows partitioning of the contributions of the parental household

Table III. Means of intrapair CF-specific FEV₁ percentile differences while living together versus living apart*

	Monozygous twins	Combined dizygous twins/siblings	T-test, P value [†] [Difference: 95%CI]
All subjects			
No.	(n = 29 pairs)	(n = 33 pairs)	
Living together (years of data)	8.0 ± 6.9 (11.4 ± 3.3)	19.2 ± 15.5 (11.3 ± 3.9)	.0006 [-11.2: -17.5, -5.0]
Living apart (years of data)	11.2 ± 11.9 (5.6 ± 5.0)	22.4 ± 16.3 (4.4 ± 3.6)	.0035 [-11.2: -18.5, -3.8]
T-test P value [‡] [Difference: 95%CI]	0.1412 [-3.3: -7.7, 1.2]	0.3091 [-3.2: -9.5, 3.1]	
ΔF508 Homozygotes only			
No.	19 pairs	17 pairs	
Living together (years of data)	6.9 ± 6.1 (11.3 ± 2.9)	21.1 ± 16.4 (12.3 ± 3.9)	.0012 [-14.2: -22.4, -6.0]
Living apart (years of data)	7.8 ± 6.5 (4.9 ± 4.2)	28.3 ± 17.4 (4.2 ± 2.7)	<.0001 [-20.5: -29.2, -11.8]
T-test P value [‡] [Difference: 95%CI]	.6630 [-0.9: -5.2, 3.4]	.1069 [-7.2: -16.1, 1.7]	

*Mean differences are the mean of the absolute differences in CF-specific FEV₁ for each twin/sibling pair for the indicated time periods. Only pairs where both members had data while living together and living apart are included.

[†]Unpaired t-test P value is for monozygous twin group versus the combined dizygous twin/sibling group.

[‡]Paired t-test P value is for living together versus living apart.

versus the effect of living with another sibling with CF. The intraindividual analysis was performed on a different subset of the study population not studied in the prior analyses and was restricted to 40 sibling pairs who had pulmonary function measures available for both siblings before and after they left the parental household.

To perform the analysis, pulmonary function was averaged for each patient when living with parents and a sibling with

CF, when living with only parents, and when living away from both sibling and parents. As shown in **Figure 4**, changes in CF-specific FEV₁ appear to be greater when patients moved away from both parents *and* siblings (ie, A to B [12.3] and C to E [12.6]), compared with moving away from only the parents (D to E [8.3]). Furthermore, the presence of another sibling with CF affected an individual's pulmonary function as demonstrated by changes in the sibling that remained with the parents as the other sibling left home (C to D [11.2]). The contribution of the sibling (C to D) and parental household (D to E) did not equal the combined total (C to E) owing to the presence of unique environmental factors throughout the period C to E.

To quantify the contribution of the parental versus affected sibling effects, intraindividual differences were placed into a regression model: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ where $Y =$ Intraindividual (absolute) difference in CF-specific FEV₁; $X_1 =$ [0 if no change, 1 if separating from a sibling with CF]; $X_2 =$ [0 if no change, 1 if leaving parental household]. As before, the intercept value (β_0) was assumed to be due to unique environmental and stochastic factors over time, after accounting for shared environmental factors because of the presence of a sibling with CF (β_1) and the parental household (β_2). The initial model demonstrated that $\beta_0 = 7.0 \pm 1.4$ (P value < .001), $\beta_1 = 4.1 \pm 1.3$ (P value = .003), and $\beta_2 = 1.3 \pm 1.2$ (P value = .310), with an overall model P value of .0110. For the final model, the nonsignificant parental variable (X_2) was dropped, yielding $\beta_0 = 8.3 \pm 1.2$ (P value < .001), and $\beta_1 = 3.7 \pm 1.3$ (P value = .007), with an overall model P value of .0072. The final intraindividual regression model generates very similar estimates for the contribution of environmental factors (unique: 8.3; shared: 3.7) as did the intrapair model (8.0; 3.2).

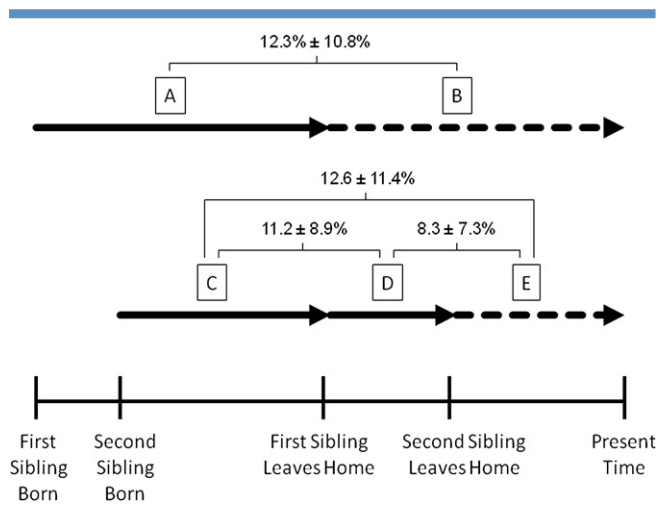


Figure 4. Differences in intra-individual pulmonary function (CF-specific FEV₁ percentiles) between different living environments. *Solid lines* represent time spent in the parental home; *dashed lines* represent time spent outside the parental home. The *lettered boxes* refer to discrete periods of time for which pulmonary function is averaged. **A**, Firstborn sibling from birth to leaving the parental home. **B**, First sibling after leaving home. **C**, Second sibling from birth to when the first sibling leaves the home. **D**, Second sibling from when the first sibling leaves home until the second sibling leaves home. **E**, Second sibling after leaving home. CF-specific FEV₁ percentiles for 40 pairs of siblings was averaged over each lettered time period, and mean absolute differences (± SD) between selected time periods are provided. Sibling pairs were included only if data existed for all lettered time periods. Siblings were not necessarily concordant for sex or born within 3 years of each other.

Discussion

Determining the causes of lung disease variation will help identify targets for ameliorating morbidity and extending life-span for individuals with CF. Our results demonstrate that environmental and genetic factors contribute approximately

equally to pulmonary function variation. As a number of studies have described shared environmental exposures that significantly affect CF pulmonary function (eg, second-hand smoke), we were surprised to find the majority of the environmental contribution is mediated through effects that act uniquely (or stochastically) on an individual. This finding was replicated using intraindividual analyses in a different study population. These similar analytic findings are illustrated in **Figure 5** when comparing monozygous twins living apart with the intraindividual analysis.

Our analyses confirm the potency of the genetic influence on pulmonary function variation^{7,8} and add several new and useful observations. Specifically, monozygous twins' pulmonary function tends to remain highly similar even after leaving home in contrast to dizygous twins and siblings. These results support ongoing efforts to identify genetic modifiers through candidate- and genome-wide approaches. Also, our results add that heritability remains relatively constant over time. This suggests either the variance caused by genetic and environmental sources remains constant or the variance from both genetic sources and the environment (or age) is increasing thus preserving heritability.²³ Decreasing correlation coefficients with age as seen in **Table IV** precludes the first explanation. At least two possibilities may underlie increasing phenotypic variation over time with constant heritability. First, gene-age interactions may exist as some modifier genes may have differing ages of onset of effect on the basis of genotype; one potential example in CF is mannose binding lectin.^{24,25} Second, cumulative environmental exposures may precipitate novel modifier gene-environment interactions. Each explanation implies that stratifying patients with CF by age should be considered to optimize searches for environmental and genetic modifiers of pulmonary function.

Although we found that most of the environmental contribution is mediated through unique factors, a closer examination of the common or shared environmental contribution turned up an unsuspected finding; namely we found that the presence of a sibling with CF results in greater intrapair differences in pulmonary function than does the parental environment. We speculate that parents caring for 2 children with CF divide their time and resources in caring for both children, but once the older sibling has left the home, the remaining sibling's pulmonary function is increased (before older sibling leaves: 62.4 ± 21.9 vs after: 70.9 ± 21.1 ; *t* test *P* value < .0001), perhaps as a result of increased resources being devoted to that remaining sibling. In contrast, leaving the parental home may induce less variation because the effects of the parental environment (eg, exercise regimens, nutritional habits) may persist even after the individual moves out. This is supported by results shown in **Figure 3** where differences in pulmonary function between siblings do not appear to increase until several years after moving out of the parental home. Alternatively, the absence of increasing differences may be due to siblings moving back into the parental home.²⁶

This study had several limitations. A component of genetic variation not included in this analysis is variation due to

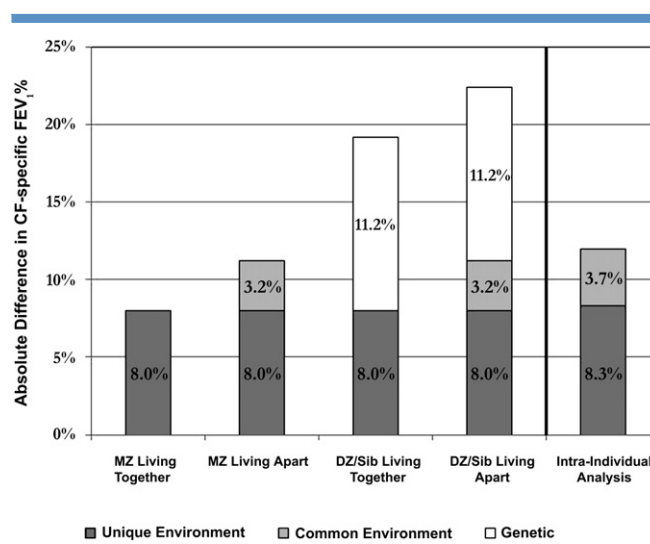


Figure 5. Summary of genetic and environmental effects on pulmonary function variation using intrapair and intraindividual differences. The data in this table are derived from the coefficients from the intrapair and intraindividual regressions in the Results section. For the 4 leftmost bars, which are based on the intra-pair regression, the sources of variation are shown for each group. The rightmost bar is based on the intraindividual regression. Of note is that the genetic contribution to pulmonary function (0.112) equals the environmental contribution (0.080 ± 0.032) for the intrapair regression. Also of note is that the estimates of the contributions of the common environment and unique environment from the intrapair analysis are very similar those generated from the independent study sub-group used in the intraindividual analysis with a different regression strategy.

CFTR because family members shared the same *CFTR* genotypes. This may not be a major shortcoming because numerous studies have indicated that *CFTR* genotype has a minimal role in variation of CF pulmonary function.³⁻⁵ Current methods of estimating sources of variation do not allow parsing of the relative contribution of unique (eg, 1 twin smoking) versus stochastic (eg, 1 twin contracts influenza) factors, or gene-environment interactions, which are likely an important source of variation, to be specifically isolated. Twins may also elect to live with each other after leaving the parental home, and this may lead to genetic effects being overestimated and environmental effects being underestimated. Our analyses of common or shared environmental factors are structured on leaving the parental home and do not necessarily capture variation caused by CF care centers or regional factors such as climate or air pollution. Finally, our data after leaving the parental home is limited to 5 years of pulmonary function data. To help address some of these limitations, the twins and siblings with CF will be reevaluated after a considerably longer duration of time living apart.

In summary, our findings demonstrate that environmental and genetic factors make contributions of approximately similar magnitude to variation in CF lung disease and justify

the need for ongoing and future studies to examine both sets of factors. Furthermore, the important contribution of unique environment (and possibly stochastic factors) will inform future searches for gene-environment interactions that modify CF pulmonary function. Identification of such interactions should provide clues into disease pathophysiological study with both a stimulus (unique exposure) and target (gene), which may lead to individualized therapies for CF and other chronic respiratory diseases. ■

We thank the CF Foundation for use of the CF Foundation Data Registry, Ase Sewall, BS, Monica Brooks, BS, the staff at the Data Registry, Nulang Wang, BS, and Sarah Ritter, BA, for CFTR genotyping, Michal Kulich, PhD, for providing conversion programs for CF-specific percentile estimates for FEV1, (served as a biostatistician for Cystic Fibrosis Foundation Therapeutic Development Network Coordinating Center), Rita McWilliams, PhD, Julie Hoover-Fong, MD, PhD, and Ada Haimosh, MD, MPH, for designing questionnaires, and most importantly the patients with CF and their families, research coordinators, nurses, and physicians who are participating in the CF Twin and Sibling Study.

Submitted for publication Nov 10, 2009; last revision received Feb 23, 2010; accepted May 7, 2010.

Reprint requests: Garry R. Cutting, MD, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Broadway Research Building Suite 551, 733 N Broadway Ave, Baltimore, MD 21205. E-mail: gcutting@jhmi.edu.

References

1. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989;245:1066-73.
2. Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science* 1989;245:1073-80.
3. Koch C, Cuppens H, Rainisio M, Madessani U, Harms H, Hodson M, et al. European Epidemiologic Registry of Cystic Fibrosis (ERCF): comparison of major disease manifestations between patients with different classes of mutations. *Pediatr Pulmonol* 2001;31:1-12.
4. Kerem E, Corey M, Kerem B-S, Rommens J, Markiewicz D, Levison H, et al. The relation between genotype and phenotype in cystic fibrosis—analysis of the most common mutation (deltaF508). *N Engl J Med* 1990;323:1517-22.
5. Cystic Fibrosis Genotype/Phenotype Consortium. Correlation between genotype and phenotype in cystic fibrosis. *N Engl J Med* 1993;329:1308-13.
6. Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry Annual Data Report 2006. 2006.
7. Vanscoy LL, Blackman SM, Collaco JM, Bowers A, Lai T, Naughton K, et al. Heritability of lung disease severity in cystic fibrosis. *Am J Respir Crit Care Med* 2007;175:1036-43.
8. Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H, Tummler B. Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. *Twin Res* 2000;3:277-93.
9. Smyth A, O’Hea U, Williams G, Smyth R, Heaf D. Passive smoking and impaired lung function in cystic fibrosis. *Arch Dis Child* 1994; 71:353-4.
10. Beydon N, Amsellem F, Bellet M, Boule M, Chaussain M, Denjean A, et al. Pulmonary function tests in preschool children with cystic fibrosis. *Am J Respir Crit Care Med* 2002;166:1099-104.
11. Campbell PW, III, Parker RA, Roberts BT, Krishnamani MR, Phillips JA, III Association of poor clinical status and heavy exposure to tobacco smoke in patients with cystic fibrosis who are homozygous for the F508 deletion. *J Pediatr* 1992;120:261-4.
12. Rubin BK. Exposure of children with cystic fibrosis to environmental tobacco smoke. *N Engl J Med* 1990;323:782-8.
13. Collaco JM, Vanscoy L, Bremer L, McDougal K, Blackman SM, Bowers A, et al. Interactions between secondhand smoke and genes that affect cystic fibrosis lung disease. *JAMA* 2008;299:417-24.
14. O’Connor GT, Quinton HB, Kneeland T, Kahn R, Lever T, Maddock J, et al. Median household income and mortality rate in cystic fibrosis. *Pediatrics* 2003;111:e333-9.
15. Schechter MS, Margolis PA. Relationship between socioeconomic status and disease severity in cystic fibrosis. *J Pediatr* 1998;132:260-4.
16. Schechter MS, Shelton BJ, Margolis PA, Fitzsimmons SC. The association of socioeconomic status with outcomes in cystic fibrosis patients in the United States. *Am J Respir Crit Care Med* 2001;163:1331-7.
17. Curtis JR, Burke W, Kassner AW, Aitken ML. Absence of health insurance is associated with decreased life expectancy in patients with cystic fibrosis. *Am J Respir Crit Care Med* 1997;155:1921-4.
18. Goss CH, Newsom SA, Schildcrout JS, Sheppard L, Kaufman JD. Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis. *Am J Respir Crit Care Med* 2004;169:816-21.
19. Wexler NS, Lorimer J, Porter J, Gomez F, Moskowitz C, Shackell E, et al. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington’s disease age of onset. *Proc Natl Acad Sci U S A* 2004; 101:3498-503.
20. Rosenstein BJ, Cutting GR. The diagnosis of cystic fibrosis: a consensus statement. *J Pediatr* 1998;132:589-95.
21. Kulich M, Rosenfeld M, Campbell J, Kronmal R, Gibson RL, Goss CH, et al. Disease-specific Reference Equations for Lung Function in Patients with Cystic Fibrosis. *Am J Respir Crit Care Med* 2005;172:885-91.
22. Introduction to Quantitative Genetics. 4th ed. Essex: Longman; 1996.
23. Falconer DS. The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Ann Hum Genet* 1967;31:1-20.
24. Dorfman R, Sandford A, Taylor C, Huang B, Frangolias D, Wang Y, et al. Complex two-gene modulation of lung disease severity in children with cystic fibrosis. *J Clin Invest* 2008;118:1040-9.
25. Muhlebach MS, MacDonald SL, Button B, Hubbard JJ, Turner ML, Boucher RC, et al. Association between mannan-binding lectin and impaired lung function in cystic fibrosis may be age-dependent. *Clin Exp Immunol* 2006;145:302-7.
26. Quittner AL, Modi A, Boyle MP. The role of independence in daily functioning for adults with cystic fibrosis: Results from the adult data for understanding lifestyle and transitions (ADULT) survey. 31 ed. 2008. p. 445.

Table I. Characteristics of included and excluded subjects*

	Group A CF Twin-Sibling Study	Group B Excluded	Group C Study population	Group D Intrapair analyses	Group E Intraindividual analyses	Group F Not used in analyses
N	1535	448	1080	406	80	594
Age as of 12/31/07 (yrs: Mean \pm SD) [†]	16.7 \pm 9.7 [0.3 – 64.6] (n = 1501)	12.8 \pm 11.4 [0.3 – 64.6] (n = 421)	18.2 \pm 8.6 [6.1 – 54.7]	18.7 \pm 8.7 [6.2 – 50.6]	26.4 \pm 6.3 [19.0 – 47.6]	16.7 \pm 8.1 [6.1 – 54.7]
Sex: %Male [‡]	51.8 (n = 1519)	48.1 (n = 439)	53.3	56.7	52.5	51.2
CFTR Genotype: % Δ F508 Homozygote [§]	49.3 (n = 1488)	46.1 (n = 408)	50.6	53.7	57.5	47.5
Maximum FEV ₁	69.3 \pm 26.3	69.2 \pm 26.0	69.3 \pm 26.3	68.3 \pm 27.8	74.8 \pm 22.1	69.2 \pm 25.7
CF% in the most recent year of data	Median: 77.3 (n = 1311)	Median: 78.1 (n = 231)	Median: 77.3	Median: 76.9	Median: 81.1	Median: 76.6

PFT, Pulmonary function test.

*Please see Figure 1 for a visual depiction of the derivation of these groups. Group A is comprised of groups B and C; group C is comprised of groups D, E, and F.

[†]Because not all individuals have PFT data available, age was calculated from a particular time rather than from the time of the most recent PFT as in Table I. Individuals in group B are younger than those in C (*t* test *P* value < .0001) because group B includes many individuals too young to perform PFTs reliably. With an ANOVA test with Bonferroni correction, group E is older than group D (*P* value < .001) and group F (*P* value < .001) because individuals in this group require both members to have pulmonary function while living together and living and apart; hence, they are older. Group D is also older than group F (ANOVA, *P* value < .001).

[‡]Groups B and C do not differ by sex frequencies (χ^2 , *P* value = .06). Groups D, E, and F also do not differ from each other by sex frequencies (χ^2 , *P* value = .23).

[§]By *CFTR* genotype frequencies, groups B and C do not differ (χ^2 , *P* value = .15), groups D and E do not differ (χ^2 , *P* value = .32), and groups D and F also do not differ (χ^2 , *P* value = .06). However, group E has a higher frequency of Δ F508 homozygotes than does group F (χ^2 , *P* value = .045). As group E was used for intraindividual analyses where an individual is compared with him or herself, we would expect *CFTR* genotype to have less of an impact on the analyses, but this possibility cannot be excluded.

^{||}Because not all individuals have PFT data for while living together or living apart, the maximum CF-specific percentile in the most recent year of PFT data was selected as an inclusive PFT measure to compare individuals in the various populations. There is no difference in this measure of pulmonary function between groups B and C (*t* test, *P* value = .99) or between groups D, E, and F (ANOVA, *P* value = .13).

Table IV. Correlation of pulmonary function (CF-specific FEV₁ percentiles) for twin and sibling pairs within discrete age groups*

Age group	Monozygous twins Mean r value	Combined dizygous twins/siblings Mean r value	Heritability estimate
6-10 y	0.6529 (n = 57 pairs)	0.4764 (n = 116 pairs)	0.35 [†]
10-14 y	0.8791 (n = 45 pairs)	0.5070 (n = 101 pairs)	0.74
14-18 y	0.9094 (n = 39 pairs)	0.6147 (n = 74 pairs)	0.59
18-22 y	0.8799 (n = 31 pairs)	0.5153 (n = 43 pairs)	0.73
22-26 y	0.8494 (n = 24 pairs)	0.3465 (n = 25 pairs)	1.00
26-30 y	0.7871 (n = 13 pairs)	0.2725 (n = 11 pairs)	1.00
30-40 y	0.5441 (n = 7 pairs)	0.3637 (n = 7 pairs)	0.36 [†]

*Derivation of r values and heritability estimates are described in the Methods section. All pairs where both members of the pair had data within the specific age range were included (out of a total of 67 monozygous twin pairs and 136 dizygous twin and sibling pairs).

[†]Lower heritability estimates observed in the 6- to 10-year-old and 30- to 40-year-old groups are likely due to increased variability in young children's pulmonary function test results and fewer numbers of older subjects, respectively.

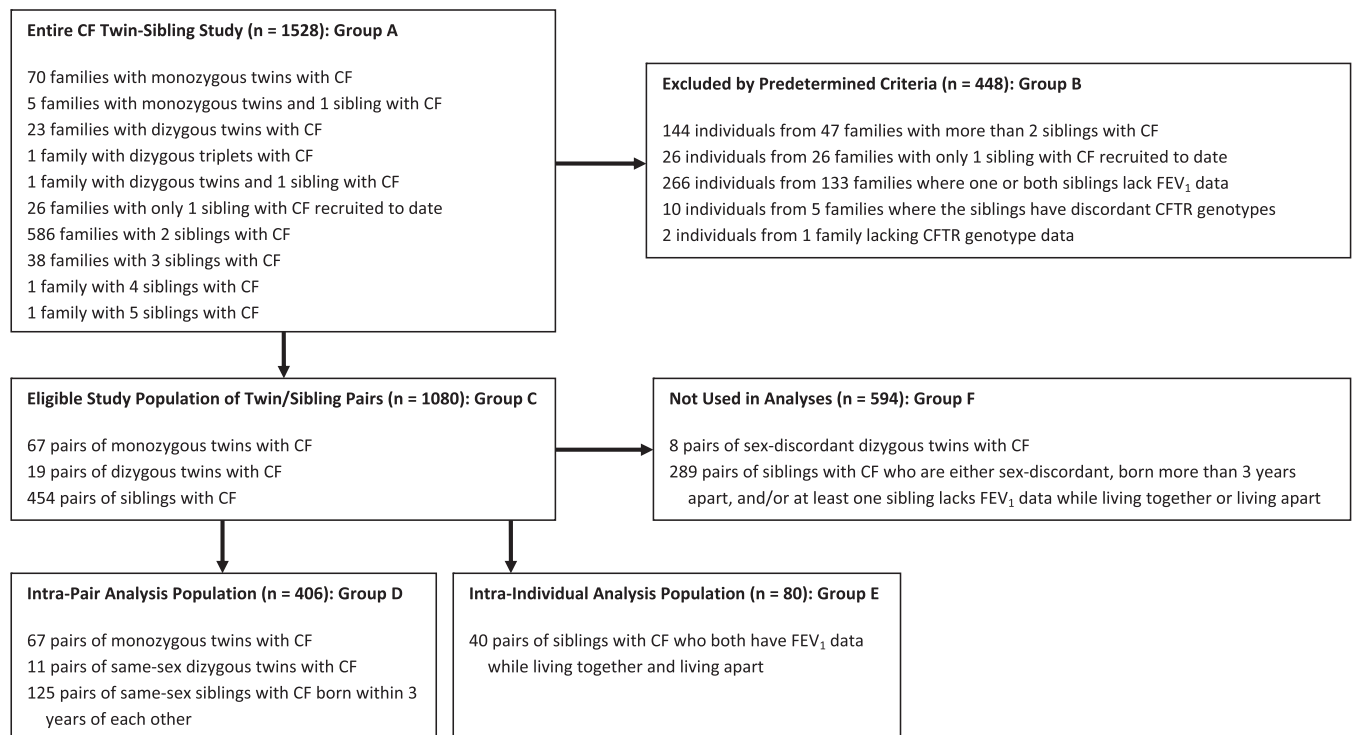


Figure 1. Study population inclusion and exclusion criteria: demographic characteristics of each group (A-F) can be found in [Table 2](#).

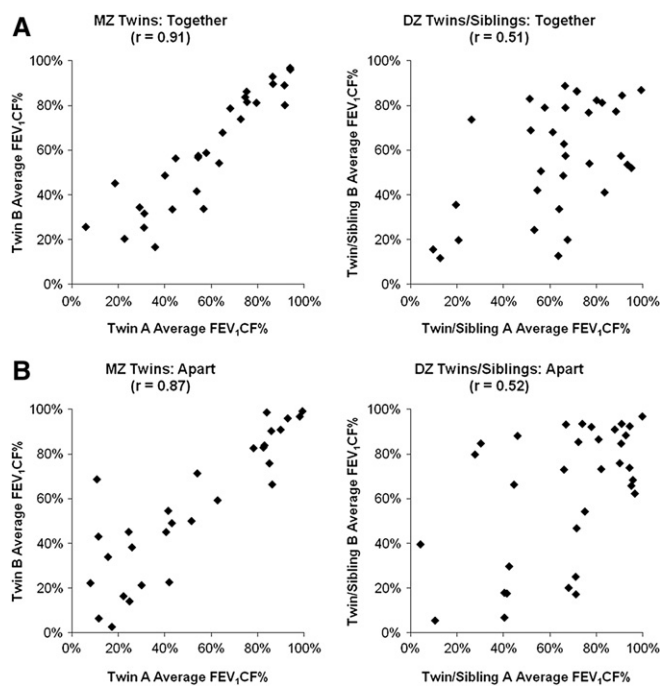


Figure 2. Dizygous twins and siblings show greater intrapair differences in pulmonary function (CF-specific FEV₁ percentiles) when living together and when living apart than monozygous twins. The plots demonstrate the concordance in average pulmonary function between the same pairs of family members when **A**, living together and **B**, living apart. Most monozygous twins within a pair either both experience improving or declining pulmonary function over time; the same is not necessarily true for dizygous twins and siblings where the members follow different trajectories.

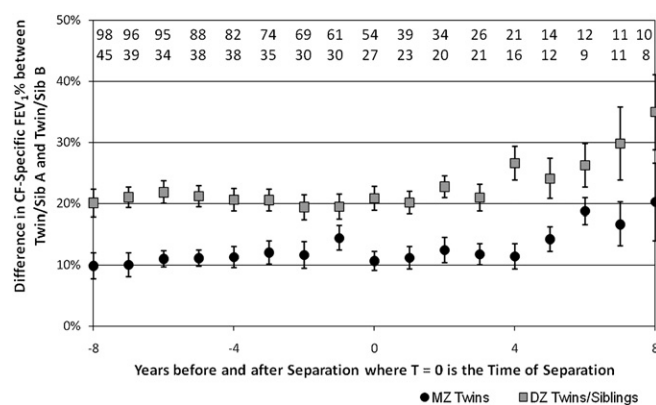


Figure 3. Mean difference in pulmonary function (CF-specific FEV₁ percentiles) increases between twins and siblings after they leave the parental home. Each point represents the mean absolute difference (\pm SE) in cross-sectional pulmonary function between family members. The numbers above the data points represent the number of pairs contributing to that time point (**Upper**, dizygous/Sib; **Lower**, monozygous). Negative years represent the time leading up to leaving the parental home, positive years the time afterward.