



Published in final edited form as:

Environ Adv. 2023 December ; 14: . doi:10.1016/j.envadv.2023.100449.

Social-environmental phenotypes of rapid cystic fibrosis lung disease progression in adolescents and young adults living in the United States

Anushka K. Palipana^{a,b}, Andrew Vancil^b, Emrah Gecili^{b,c}, Erika Rasnick^b, Daniel Ehrlich^{a,b}, Teresa Pestian^b, Eleni-Rosalina Andrinopoulou^{d,e}, Pedro M. Afonso^{d,e}, Ruth H. Keogh^f, Yizhao Ni^g, Judith W. Dexheimer^{c,h}, John P. Clancyⁱ, Patrick Ryan^{b,c}, Cole Brokamp^{b,c}, Rhonda D. Szczesniak^{b,c,j,*}

^aDuke University, Durham, NC, United States

^bDivision of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

^cDepartment of Pediatrics, University of Cincinnati, Cincinnati, OH, United States

^dDepartment of Biostatistics, Erasmus Medical Center, Rotterdam, the Netherlands

^eDepartment of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands

^fLondon School of Hygiene and Tropical Medicine, London, UK

^gKaiser Permanente, Denver, CO, United States

^hDivision of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

ⁱCystic Fibrosis Foundation, Bethesda, MD, United States

^jDivision of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

Abstract

Background: Cystic fibrosis (CF) is a genetic disease but is greatly impacted by non-genetic (social/environmental and stochastic) influences. Some people with CF experience rapid decline, a precipitous drop in lung function relative to patient- and/or center-level norms. Those who

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author at: Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave (MLC 5041), Cincinnati, OH 45229, United States. Rhonda.Szczesniak@cchmc.org (R.D. Szczesniak).

Contributions

The authors thank the Cystic Fibrosis Foundation for use of Patient Registry (CFFPR) data to conduct this study and the patients, care providers, and clinic coordinators at U.S. CF centers for their contributions to the CFFPR. Requests to access the CFFPR analysis data utilized for this study may be sent to the Foundation for review/approval: data-requests@cff.org.

Declaration of Competing Interest

Author RDS serves on the Cystic Fibrosis Foundation Patient Registry Committee. The remaining authors have no conflicts of interest to report.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.envadv.2023.100449.

experience rapid decline in early adulthood, compared to adolescence, typically exhibit less severe clinical disease but greater loss of lung function. The extent to which timing and degree of rapid decline are informed by social and environmental determinants of health (geomarkers) is unknown.

Methods: A longitudinal cohort study was performed (24,228 patients, aged 6–21 years) using the U.S. CF Foundation Patient Registry. Geomarkers at the ZIP Code Tabulation Area level measured air pollution/respiratory hazards, greenspace, crime, and socioeconomic deprivation. A composite score quantifying social-environmental adversity was created and used in covariate-adjusted functional principal component analysis, which was applied to cluster longitudinal lung function trajectories.

Results: Social-environmental phenotyping yielded three primary phenotypes that corresponded to early, middle, and late timing of peak decline in lung function over age. Geographic differences were related to distinct cultural and socioeconomic regions. Extent of peak decline, estimated as forced expiratory volume in 1 s of % predicted/year, ranged from 2.8 to 4.1 % predicted/year depending on social-environmental adversity. Middle decliners with increased social-environmental adversity experienced rapid decline 14.2 months earlier than their counterparts with lower social-environmental adversity, while timing was similar within other phenotypes. Early and middle decliners experienced mortality peaks during early adolescence and adulthood, respectively.

Conclusion: While early decliners had the most severe CF lung disease, middle and late decliners lost more lung function. Higher social-environmental adversity associated with increased risk of rapid decline and mortality during young adulthood among middle decliners. This sub-phenotype may benefit from enhanced lung-function monitoring and personalized secondary environmental health interventions to mitigate chemical and non-chemical stressors.

Keywords

Chemical stressors; Cluster analysis; Community deprivation; Environmental health epidemiology; Lung function; Medical monitoring

1. Introduction

Cystic fibrosis (CF) is a rare genetic disease, but nearly 50 % of the variation in lung function is attributable to non-genetic (social, environmental and stochastic) factors (Collaco et al., 2010). The crucial role of social and environmental exposures in the development and exacerbation of lung diseases such as asthma and chronic obstructive pulmonary disease (COPD) is widely acknowledged (Guarnieri and Balmes, 2014; Hansel et al., 2016; Louisias et al., 2019). Air pollutants can harm lung tissue directly upon exposure and indirectly by producing reactive oxygen species and causing systemic inflammation. Despite the well-established link between air pollution and respiratory disorders, the impact of social and environmental factors on people living with CF remains largely unexplored. Furthermore, the effects of social and environmental exposures on CF lung function decline have primarily been studied one exposure domain at a time. Examples of two separate but frequently studied exposures are air pollution (Blayac et al., 2022) and Medicaid insurance use (Schechter et al., 2001). While the former has been linked primarily to pulmonary

exacerbation onset (Goeminne et al., 2013; Goss et al., 2004), which is largely driven by acute drops in lung function, the latter has often corresponded to prolonged drops in lung function (Schechter and Margolis, 1998). Recent tobacco smoke exposure studies using pediatric data from the U.S. Cystic Fibrosis Foundation Patient Registry (CFFPR) found that first- or second-hand smoking cessation associated with improved lung function over time (Oates et al., 2021), and, for individuals treated with tezacaftor/ivacaftor, there was dampened benefit from modulator therapy shown by continued divergence in lung function trajectories of exposed and unexposed groups (Baker et al., 2021).

The CF lung disease process is typically characterized by acute and prolonged drops in lung function. Attenuated decreases in lung function relative to patient- and/or center-level norms, clinically termed rapid decline, typically manifest during adolescence and early adulthood (Konstan et al., 2007; Vandenbranden et al., 2012). Early rapid decline tends to happen in individuals with below-average lung function, infections, poor/deteriorating nutrition, increased antibiotic use, and more frequent hospitalizations. Meanwhile, regardless of clinical phenotype, lung function trajectories exhibit nonlinear, heterogeneous patterns over this age range (Harun et al., 2016; Szczesniak et al., 2013; Vandenbranden et al., 2012). A classification study of F508del homozygotes indicated that lung function (initial level and rate of decline) coupled with survival percentiles at 20 years of age accurately distinguished mild and severe pulmonary phenotypes from the Gene Modifier Study (Schluchter et al., 2006). An effort to phenotype rapid decline was made using the aforementioned CFFPR data source while accounting for both age-related nonlinearity and heterogeneity in lung function trajectories suggested that the magnitude of peak lung function decline is similar; however, timing of peak decline differed according to early, middle, and late ages during adolescence/early adulthood: 12.9, 16.3, and 18.5 years, respectively (Szczesniak et al., 2017). In addition to the above risk factors of rapid decline, the early phenotype was associated with Medicaid insurance use. A single-center study including young adults with CF developed an index score of rapid decline using quantiles of lung function trajectories, showing that nutrition- and sex-related effects were more influential in higher quantiles corresponding to worsening disease severity (Denaro et al., 2020).

Past approaches have illuminated subsets of social and environmental exposures that are correlates of rapid decline, primarily studied independent from phenotyping. As a result, social-environmental phenotypes of rapid decline have not been characterized. In this study, we hypothesized that a comprehensive evaluation of social and environmental exposures (geomarkers), including the formation of an adversity index, would yield distinct age-related phenotypes of rapid decline. Additionally, we hypothesized that more severe social-environmental phenotype scores would associate with worsening patient characteristics that are routinely surveilled as part of CF care, including genotype, lung infections, smoking status, and comorbidities.

2. Methods

2.1. Study population and health data

Inclusion criteria: We performed a longitudinal cohort study of patients aged 6–21 years followed in the CFFPR (over the timeframe 1997–2017) approved by the local Institutional

Review Board (Protocol ID: 2018-4839). The CFFPR has been used to track demographic and clinical characteristics of people living with CF in the U.S. for many decades with a recently estimated coverage of 77 % allowing for generalizability across the population (Cromwell et al., 2023). The primary outcome was forced expiratory volume in 1 s (FEV₁) of % predicted from Global Lung Function Initiative reference equations (Quanjer et al., 2012). Age (in years) was the time variable. Individuals with less than 7 quarterly FEV₁ observations over the study timeframe or who never reported ZIP Code were excluded. Post-lung transplantation data was censored. Baseline was individually defined as first observed FEV₁ during the study period once the patient was aged ≥ 6 years.

Routinely collected health measures: Non-time-varying CFFPR variables were genotype defined by F508del alleles (homozygous, heterozygous, or neither/unknown), race (White or non-White), ethnicity (Hispanic or non-Hispanic), sex (male or female), age at CF diagnosis, any recorded use of pancreatic enzymes (yes/no). Time-varying variables were private insurance use (yes or no); lung infections (each coded as yes or no): *Pseudomonas aeruginosa* (Pa) and Methicillin-resistant *Staphylococcus aureus* (MRSA); CF-related diabetes (CFRD) status; modulator use: ivacaftor or lumacaftor/ivacaftor; reporting either first- or secondhand tobacco smoke exposure (obtained from annual self-report). Pulmonary exacerbation (PEX) frequency was defined by counting the number of events with intravenous antibiotic use that required hospitalization within the prior year.

2.2. Geomarker assessments

Social and environmental exposures (geomarkers) were linked to ZIP Codes in the CFFPR by using ZIP Code tabulation areas (ZCTAs, a census-derived geography for 5-digit residential ZIP Codes). Geomarkers (respective data sources) included traffic proximity, ozone and PM_{2.5} concentrations, diesel particulate matter, and respiratory hazard index (Environmental Protection Agency Environmental Justice Screen Index (United States Environmental Protection Agency, 2015), landcover (% of greenspace, impervious, and tree canopy areas derived from the National Land Cover Database (Jin et al., 2019)), total crime (Applied Geographic Solutions (Federal Bureau of, 2018)), neighborhood material deprivation index (Brokamp et al., 2019) (assesses extent of poverty, vacant housing, assisted income, education level, median income and health insurance coverage for a given neighborhood), lengths and densities of primary and secondary roads (derived from geospatial data in the Topologically Integrated Geographic Encoding and Referencing system (United States Department of Commerce, 2018)). Rationale for including each geomarker was based on components of a postulated CF disease-outcome model (Schechter, 2011) and their roles as potential correlates of rapid CF lung disease progression from more recent literature review (Szczesniak et al., 2020) (e-Table 1). We imputed missing data on demographic/clinical variables and ZIP Codes based on the nature of collection for each CFFPR variable (e-Table 2). Backward-imputing was employed for geomarkers and ZIP Codes that were not reported until after baseline.

2.3. Statistical analyses

We summarized all variables at baseline and over follow-up as mean (standard deviation or SD) for continuous variables and n (%) for categorical variables. Analyses were

implemented using R and MATLAB (code and packages described in e-Section A). We performed a two-stage approach to obtain social-environmental phenotypes of rapid decline. In the first stage, principal components analysis (PCA) was used to characterize relationships between geomarkers at baseline and acquire a composite geomarkers score for subsequent cluster analysis. For the second stage, we performed covariate-adjusted functional principal components analysis (FPCA) for longitudinal data (Jiang and Wang, 2010), using mean quarterly FEV₁% predicted as the outcome variable.

Since the available FPCA method only accommodates one covariate aside from the time variable (which is age in our approach), we included the aforementioned social-environmental adversity index as the covariate. Detailed implementation and results, including selecting the first functional principal component and forming tertiles to represent phenotypes of rapid decline according to age, are provided in e-Section B.

After performing the two-stage analysis, we examined associations between derived phenotypes with respect to individual demographic/clinical characteristics using Wilcoxon rank sum with Bonferroni-adjusted p-values and Chi-square tests. A p-value less than 0.05 was considered statistically significant.

2.4. Sensitivity analyses

We assessed the impact of identifying phenotypes using average rather than baseline values of the geomarkers in PCA (first stage). Second-stage sensitivity analyses evaluated (i) potential selection bias from requiring a minimum number of FEV₁ measurements; (ii) how FPCA findings may be impacted by variable length of follow-up between patients; (iii) impact of modulator initiation; (iv) impact of loss-to-follow-up in the CFFPR due to 2003–2006 intake changes, which included introduction of a web-based platform to capture encounter-level data in 2003 followed by detailed medication collection that began in 2006 (Knapp et al., 2016); (v) extent to which imputation of later-observed geomarkers into earlier time periods (largely prior to 2015) affected results.

3. Results

3.1. Study population and routinely monitored characteristics

There were 24,228 individuals with 664,267 quarterly measurements who met inclusion criteria (e-Fig. 1). The analysis cohort primarily included F508del homozygotes, had slightly more males, and many did not have private insurance during the analysis timeframe (Table 1 – Overall cohort). Infections and CFRD prevalence increased expectedly over follow up. Very few individuals used modulators at baseline, but prevalence increased over follow up. Reported tobacco smoke also increased over time. Most individuals did not undergo lung transplant and remained alive through follow up. Geomarker data was variable across the overall cohort (Table 2). Deprivation index, which ranges from 0 to 1 with higher values being associated with higher levels of material community deprivation, for the overall CF cohort was below the national average, estimated to be 0.37 and 0.35 (computed by weighting each tract-level deprivation index by its population under age 18) for 2015 and 2018, respectively.

The first PC score was retained from scree-plotting (e-Fig. 3) and used to collectively measure extent of social-environmental adversity (higher scores implied more negative social and/or environmental exposure levels; details shown in e-Section A). Social-environmental adversity was primarily driven by crime, landcover, traffic-related air pollution, and density and length of primary and secondary roadways, respectively (loadings for rotated PC₁, e-Table 6). Although not used to form the social-environmental adversity index, the second principal component mainly linked to deprivation index, ozone concentration, primary roadway length, and secondary roadway density (loadings for rotated PC₂, e-Table 6).

3.2. Social-environmental lung phenotypes over age

Functional principal component analysis (FPCA) suggested highly heterogeneous lung function decline according to age and social and environmental exposures, although all individuals lost some degree of lung function by early adulthood (Fig. 1A). Overall, increased social-environmental adversity corresponded to higher lung function at younger ages, but the degree of rapid decline was more pronounced in adolescence and early adulthood, indicating declines of more than 4 % predicted/year for some individuals (Fig. 1B). Upon segmenting these unusual results, subjects with less social-environmental adversity also experienced lung function decline, but at a slower rate, compared to counterparts with high social-environmental adversity (e-Fig. 4). Specifically, excluding individuals with outlier values for the social-environmental adversity index showed that higher social-environmental adversity associated with more rapid lung function decline over age (e-Fig. 5).

We classified individuals into three distinct social-environmental phenotypes of rapid decline over age (early, middle, or late) based on FPC₁ first and third quantiles from FPCA cluster analysis (e-Table 5). Further segmenting early, middle, and late decliners into subgroups according to median social-environmental adversity, there was within-phenotype variability (Fig. 2). Early decliners experienced similar timing and extent of peak decline regardless of social-environmental adversity (Fig. 2A versus 2D). Middle decliners had similar peak decline, but those with greater extent of social-environmental adversity experienced peak decline an average of 14.4 months or 1.2 years earlier than those with lower social-environmental adversity (Fig. 2B versus 2E). Late decliners with lower social-environmental adversity had slightly earlier peak decline (approximately four months earlier), compared to their counterparts with higher social-environmental adversity (Fig. 2C versus 2F).

Specific geomarkers primarily drove differences in social-environmental adversity among the early, middle, and late decline phenotypes (Table 2). The earlier declining phenotypes tended to reside in areas with higher community deprivation and crime. The average deprivation index for late decliners was lower than the aforementioned national averages, while average deprivation for early decliners was similar to national averages. Compared to late decliners, early decliners tended to reside in areas with slightly higher green space and imperviousness. Early decliners also tended to live in areas with higher air pollution, measured by the respiratory hazard index, ozone concentration, proximity to traffic, and

diesel particulate matter. Middle decliners tended to live closer to longer/denser secondary roadways compared to late decliners. While early decliners tended to live closer to longer secondary roadways compared to late decliners, they also live closer to denser secondary roadways compared to middle decliners.

Estimated peak decline and degree of social-environmental adversity were jointly mapped across the contiguous U.S. (Fig. 3). While findings were heterogeneous, the most severe declines coupled with greatest social-environmental adversity (purple shaded areas) were concentrated in western and southern regions. These regions also included areas with higher social-environmental adversity but less rapid decline (pink shaded areas). In examining designations from the American Communities Project, areas with the highest social-environmental adversity and peak decline levels corresponded to Hispanic Centers, Working Class Country (rural communities) and Big City (densely populated urban areas). Social-environmental adversity appeared lowest in the Central Plains, but these areas were marked with higher peak declines (green-shaded areas). A subset of the map was not estimable (black shaded areas). The map referred to as Fig. 3 has been made available in an interactive R Shiny app. This app allows readers to explore the joint variation of estimated peak decline and the degree of social-environmental adversity on a state-by-state basis. This app can be accessed using the following link: https://medmonitoring.shinyapps.io/Interactive_map3digZip/

3.3. Clinical subgrouping of phenotypes

Preliminary comparisons between early, middle, and late decline phenotypes (Table 1) showed that middle decliners had higher representation of F508del homozygotes and males, compared to other phenotypes. Middle decliners were slightly older with substantially lower baseline lung function. Having no insurance or use of public insurance, lung infections and CFRD diagnosis were more prevalent in early decliners. However, late decliners had the highest prevalence of reported impaired glucose tolerance. Middle decliners had the highest reported use of pancreatic enzymes. Ivacaftor use was most common among late decliners, followed by middle, then early decliners. Middle decliners had the highest reported use of lumacaftor/ivacaftor, followed by late, then early decliners. Reported tobacco smoke exposure was most prevalent among middle decliners. Early decliners had the lowest PEx frequency but highest rates of hospitalizations prior to baseline and lung transplant and death over follow-up. Kaplan–Meier analysis of phenotype-specific survival probabilities extending beyond early adulthood suggest mortality peaks occurred in early adolescence and again in early adulthood, followed by exceedingly higher rates of lung transplant/death in the early phenotype (e-Fig. 2).

3.4. Sensitivity analyses

The social-environmental adversity index was similar when performing PCA on average geospatial values, compared to using baseline geospatial values (e-Table 6). When comparing the included individuals in the analysis cohort to those who were excluded due to an insufficient number of FEV₁ measurements, we found that the excluded cohort had higher prevalence of F508del homozygotes, and those with no completed genotype were older, on average, and had slightly higher rates of Pa infection and modulator use

(e-Table 7). However, distributions of FPC_1 scores were similar (e-Table 8). Variable length of follow-up between patients did not apparently impact FPCA results based on correlation analyses (e-Fig. 7). We further examined how rate of follow up (number of visits/year) varied geographically and found that urban and rural areas had similar median years of follow-up but differing ranges: 5.67 (1.52-18.02) versus 5.56 (1.34-27.13). To consider how loss to follow up may have been impacted by CFFPR-related intake changes during 2003–2006, we summarized social-environmental adversity and degree of rapid decline separately for those individuals entering before versus after 2006. Distributions of scores from FPCA and geographical associations were similar to primary analysis (e-Fig. 10), indicating that results were not impacted by data intake changes. Detailed results are provided in e-Section C. We evaluated impact of geospatial imputation by comparing distributions of FPC_1 scores from individuals with observations only from 2015 or later to the overall cohort and to those who only had observations prior to 2015 (e-Fig. 11). Scores were largely similar among these groups; those observed 2015 and onward had a slightly higher median score.

3.5. Converting phenotypes into digestible clinical information

While the results have different implications among the three main social-environmental phenotypes, we found that existing strategies for maintaining lung function in CF under each phenotype could be layered with enhanced social and environmental health interventions (Table 3), which are based on results of this study and previously published literature (Khreis et al., 2023; McArdle et al., 2023; McGarry et al., 2017; Morgan et al., 2013; Oates and Schechter, 2016; Park et al., 2021). Routine or more frequent social needs screening, for example, could be made available for people with CF who live in communities with high deprivation or crime levels. We found that the nature of potential environmental health interventions ranges from personalized, such as installing HEPA filters in the home, to policy-level alerts or actions, including smog notifications during periods in which air pollution spikes are present or anticipated due to extreme climate events. We observed higher within-area variation compared to between areas with respect to social-environmental adversity and risk of rapid decline (e-Fig. 12), implying that people with CF and their families who make long-distance moves to different geographic regions (e.g., moving from the Midwest to the Eastern seaboard) may be subject to the same risk of rapid decline if social-environmental adversity is similar between regions. However, relocating within the same city but to a community with less social-environmental adversity may result in lower risk of rapid decline; although such a measure may be unrealistic for many with CF.

Another clinically informative takeaway for individual patients, providers and researchers may be the review of each social and environmental exposure based on how pronounced it is for their given residential area or where they will spend a significant amount of time (e-Fig. 13), in light of their individual clinical risk factors (Table 3). From the patient and clinician perspective, this more granular review could identify risk of environmental triggers of PEx, such as seasonal or event-specific spikes in air pollution levels that could be mitigated with masking. Masking strategies could be developed and implemented, similar to those being undertaken by some people with CF and their families during peak influenza season. For health equity researchers, these data may pinpoint communities suffering outcome disparities and enable clinics to implement enhanced screening of psychosocial needs.

4. Discussion

This longitudinal cohort study of children and young adults living with CF in the U.S. shows that both social and environmental stressors associate with greater degree and earlier timing of rapid lung function decline, and that, to some extent, social-environmental phenotypes of lung function decline are associated with a subset of routinely collected demographic and clinical characteristics. We observed that young people with early decline tend to reside in areas with higher total crime rates and community deprivation; in contrast, their late-declining counterparts typically reside in low-crime, low-deprivation areas, which highlights the significant role of non-chemical, socioeconomic stressors on CF disease. We also identified geographic regions with elevated ambient air pollution (traffic proximity and roadway length/density) and less greenspace in which young people with CF may be at the highest risk of early, rapid lung function decline, but we found that these areas represent heterogeneous environmental conditions spread across the US. Social-environmental adversity in CF was heavily influenced by levels of chemical stressors, such as those from traffic-related air pollution exposure. This finding corroborates a recent, pediatric single-center study performed in the Midwest, which identified elemental carbon attributable to traffic as a strong predictor of rapid CF lung function decline; while not statistically significant, greenspace and community deprivation were also selected in the final prediction model (Gecili et al., 2023). In the current study, community deprivation, which is a non-chemical stressor, in the CF population was estimated to be beneath the national average. To provide some context, given that the population level SD in 2018 was 0.14, the difference of 0.02 units in the deprivation index is roughly 14 % of the SD.

The current study corroborates earlier work suggesting that individuals with the highest levels of lung function are late decliners but cumulatively lose more lung function over childhood and adolescence than their counterparts who maintain lower lung function (Vandenbranden et al., 2012). From characteristics of late decliners identified in the current study, an outlying percentage maintained high lung function initially in the presence of extreme social-environmental adversity, while the majority of this phenotype had higher social-environmental adversity associated with more rapid decline. Identification of three prominent clusters reflects prior joint longitudinal-survival modeling of the CFFPR that also identified three distinct latent classes of FEV₁ progression (Andrinopoulou et al., 2020). Mortality among phenotypes diverged at early adolescence for those with the most severe declines, while middle- and late-declining phenotypes had similar survival probabilities. These mortality peaks are consistent with prior literature (Vandenbranden et al., 2012) but also highlight the risk to a newly identified middle-decliner subgroup susceptible to early mortality and high social-environmental adversity. It is possible that the influence of social-environmental factors may be less pronounced for individuals who decline early or late compared to their middle declining counterparts. Additionally, the impact of social-environmental adversity on lung function may vary depending on the timing of exposure and age could serve as a confounding variable in the clustering of environmental factors and rapid decline. Tailored secondary prevention strategies (e.g.: avoiding living near highways or high pollution areas; use of HEPA filters in the home (James et al., 2020)), could benefit this sub-phenotype the most. How social-environmental adversity mediates effectiveness of

established strategies, (e.g., more frequent monitoring and timely treatment of lung function declines (Schechter et al., 2018)) requires additional research.

Our work substantiates recent findings that chemical stressors (e.g., tobacco smoke exposure) have explanatory value distinct from socioeconomic deprivation (Oates et al., 2020). In COPD, the relationship between ambient air pollution and lung function appears to be moderated by sex, household income, and occupation type (Doiron et al., 2019). Specific chemical stressors have not been widely studied in CF, but causal pathways identified in asthma link air pollution to exacerbation (Pfeffer et al., 2021). Clinical manifestations among CF, asthma and COPD are distinctive (e.g.: given the prolonged versus acute nature of exacerbation onset in CF versus asthma, or the nature of airway obstruction in CF versus COPD), suggesting that causal pathways between chemical stressors and disease severity/onset differ.

The current study utilized mean quarterly FEV₁% predicted measurements, while other CF epidemiological studies have utilized different aggregates, such as the maximum. Recent advances in linear mixed effects models, which have been instrumental in monitoring and predicting the natural history of CF lung disease, suggest that use of mean or maximum quarterly FEV₁% predicted measurements result in similar trajectory estimates (Szczesniak et al., 2023). Functional data analysis techniques such as FPCA can be applied to temporally collected FEV₁% predicted measurements to identify longitudinal functional data with similar characteristics and cluster them together. Linear mixed effects models, which estimate associations between the longitudinal outcome variable and a set of explanatory variables, have led to a consensus that rapid decline often occurs between adolescence and early adulthood, although the specific timing varies across studies. In this study, using FPCA on CF and social and environmental exposure data not only helped identify patients with phenotypes at risk of rapid lung function decline but also characterized the timing of rapid lung function decline within each phenotype. This approach can enhance opportunities for more timely and targeted interventions for patients clustered into high-risk phenotypes.

This study has inherent limitations, including those previously described with CFFPR analyses (Schechter, 2008) and the large extent of missing data in income and education variables (Cystic Fibrosis Foundation, 2022), which precluded their use in the current study as a measure of individual-level socioeconomic status. There are differing amounts of overlap among the geospatial and clinical data sources, which make it difficult to produce a single, contemporaneous collection of data. Due to computational issues with FPCA of data with short-term follow-up, we were unable to restrict follow-up to the more contemporaneous time period in which select geospatial markers were observed (2015 and up). To offer some insight into how the findings are impacted by contemporaneity, we conducted several sensitivity analyses, but acknowledge that the findings are still based on the original FPCA with assumed imputations of earlier data and therefore subject to bias especially over earlier time periods of follow-up. There is also the influence of bigger data sets on precision estimates and statistical significance (Cox et al., 2018). Variability in social-environmental adversity and rapid decline was observed in aggregated estimates throughout the US, but ZIP-Code-specific inference was intentionally limited for patient privacy. U.S. Postal Service ZIP Codes are designed to facilitate the distribution of mail,

which makes the resulting ZCTAs and geospatial resolution dependent upon residential size (i.e.: cities typically have far more ZIP Codes than rural areas). Opposing PCA loadings on roadway lengths and densities show the impact of the area of ZIP Codes. An example of the granularity impact is provided showing secondary roadway densities and aggregation for a given city and rural areas (e-Fig. 14). Covariate-adjusted FPCA only allows for a single covariate, which necessitated a two-stage approach, but the resulting social-environmental adversity index was partially driven by previously identified features, such as air pollution (Goss et al., 2004), while examining novel geospatial markers. This index could enhance research on CF-specific social and environmental determinants of health by aggregating exposures across a breadth of young people with CF living in the US. Covariate adjustment with geospatial markers appeared to improve upon marginal FPCA in which identified phenotypes exhibited stronger differential selection according to baseline age (Szczeniak et al., 2017). Air temperature and seasonality were not considered, given that these characteristics typically vary over time for many locations. However, geospatial markers of traffic-related air pollution, which can be driven by increased temperatures within warmer seasons, were included. Furthermore, sensitivity analysis suggested that incorporating time-varying exposures may not lead to differing conclusions. As the study focused on a specific timeframe, patients receiving treatment with the highly effective CFTR modulator, elexacaftor-tezacaftor-ivacaftor (ETI), were not included in the analysis. However, it is possible to make assumptions about the impact of ETI treatment on the phenotypes described during the ETI era, such as the potential attenuation of lung function decline; however, effects may be dampened depending on social-environmental influences on extent to which CFTR is modulated. Considering the available literature, it is plausible that the average trajectory levels will rise but with a variable degree of rate of decline (Nichols et al., 2021).

5. Conclusion

Cystic fibrosis lung phenotypes are characterized by social-environmental adversity and a subset of routinely monitored demographic and clinical characteristics. Middle decliners have the greatest differences in timing of peak decline related to social-environmental adversity and mortality. The present study offers a comprehensive profile of geographic exposure risks and how individuals with mild to moderate amounts of lung disease who are subject to adverse social-environmental exposures may maximally benefit from personalized (rather than primary or global) environmental health interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

Supported by grants from the National Institutes of Health (R01HL141286 and P30DK117467) and the Cystic Fibrosis Foundation (GECIL20F0 and Naren19R0). The sponsors had no role in study design or manuscript preparation. A pre-submission copy of the manuscript was provided to the Cystic Fibrosis Foundation Patient Registry Committee in accordance with data request/use policies.

Data availability

The authors do not have permission to share data.

Abbreviations:

CF	cystic fibrosis
CFFPR	Cystic Fibrosis Foundation Patient Registry
CFRD	cystic fibrosis related diabetes
CI	confidence interval
FEV₁	forced expiratory volume in 1 s
FPC₁	first functional principal component
FPCA	functional principal components analysis
MRSA	Methicillin-resistant <i>Staphylococcus aureus</i>
OR	odds ratio
Pa	<i>Pseudomonas aeruginosa</i>
PC	principal component
PCA	principal components analysis
PEx	pulmonary exacerbation
Q1	first quartile
Q3	third quartile
SD	standard deviation
ZCTA	ZIP Code tabulated area.

References

- Andrinopoulou ER, Nasserinejad K, Szczesniak R, Rizopoulos D, 2020. Integrating latent classes in the Bayesian shared parameter joint model of longitudinal and survival outcomes. *Stat. Methods Med. Res* 29 (11), 3294–3307. [PubMed: 32438854]
- Baker E, Harris WT, Rowe SM, Rutland SB, Oates GR, 2021. Tobacco smoke exposure limits the therapeutic benefit of tezacaftor/ivacaftor in pediatric patients with cystic fibrosis. *J. Cyst. Fibros* 20 (4), 612–617. [PubMed: 33023836]
- Blayac M, Coll P, Urbach V, Fanen P, Epaud R, Lanone S, 2022. The impact of air pollution on the course of cystic fibrosis: a review. *Front. Physiol* 13, 908230. [PubMed: 35721541]
- Brokamp C, Beck AF, Goyal NK, Ryan P, Greenberg JM, Hall ES, 2019. Material community deprivation and hospital utilization during the first year of life: an urban population-based cohort study. *Ann. Epidemiol* 30, 37–43. [PubMed: 30563729]

- Collaco JM, Blackman SM, McGready J, Naughton KM, Cutting GR, 2010. Quantification of the relative contribution of environmental and genetic factors to variation in cystic fibrosis lung function. *J. Pediatr* 157 (5), 802–807 e801-803. [PubMed: 20580019]
- Cox DR, Kartsonaki C, Keogh RH, 2018. Big data: some statistical issues. *Stat. Probab. Lett* 136, 111–115. [PubMed: 29899584]
- Cromwell EA, Ostrenga JS, Todd JV, Elbert A, Brown AW, Faro A, Goss CH, Marshall BC, 2023. Cystic fibrosis prevalence in the United States and participation in the Cystic Fibrosis Foundation Patient Registry in 2020. *J. Cyst. Fibros* 22 (3), 436–442. [PubMed: 36922288]
- Cystic Fibrosis Foundation, (2022). Cystic Fibrosis Foundation Patient Registry: 2021 Annual Data Report. Retrieved 13 November 2023 from www.cff.org/medical-professionals/patient-registry.
- Denaro K, Bailey B, Conrad DJ, 2020. Quantifying disease severity of cystic fibrosis using quantile regression methods. *J. Data Sci* 18 (1), 148–160.
- Doiron D, de Hoogh K, Probst-Hensch N, Fortier I, Cai Y, De Matteis S, Hansell AL, 2019. Air pollution, lung function and COPD: results from the population-based UK Biobank study. *Eur. Respir. J* 54 (1).
- Federal Bureau of Investigation (2018). Crime Data Explorer, 2014-2017, Most Current Estimates and Projections. Retrieved 15 December 2022 from <https://cde.ucr.cjis.gov>.
- Gecili E, Brokamp C, Rasnick E, Afonso PM, Andrinopoulou ER, Dexheimer JW, Clancy JP, Keogh RH, Ni Y, Palipana A, Pestian T, Vancil A, Zhou GC, Su W, Siracusa C, Ryan P, Szczesniak RD, 2023. Built environment factors predictive of early rapid lung function decline in cystic fibrosis. *Pediatr. Pulmonol* 58 (5), 1501–1513. [PubMed: 36775890]
- Goeminne PC, Kicinski M, Vermeulen F, Fierens F, De Boeck K, Nemery B, Nawrot TS, Dupont LJ, 2013. Impact of air pollution on cystic fibrosis pulmonary exacerbations: a case-crossover analysis. *Chest* 143 (4), 946–954. [PubMed: 23081770]
- Goss CH, Newsom SA, Schildcrout JS, Sheppard L, Kaufman JD, 2004. Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis. *Am. J. Respir. Crit. Care Med* 169 (7), 816–821. [PubMed: 14718248]
- Guarnieri M, Balmes JR, 2014. Outdoor air pollution and asthma. *Lancet* 383 (9928), 1581–1592. [PubMed: 24792855]
- Hansel NN, McCormack MC, Kim V, 2016. The effects of air pollution and temperature on COPD. *COPD* 13 (3), 372–379. [PubMed: 26683097]
- Harun SN, Wainwright C, Klein K, Hennig S, 2016. A systematic review of studies examining the rate of lung function decline in patients with cystic fibrosis. *Paediatr. Respir. Rev* 20, 55–66. [PubMed: 27259460]
- James C, Bernstein DI, Cox J, Ryan P, Wolfe C, Jandarov R, Newman N, Indugula R, Reponen T, 2020. HEPA filtration improves asthma control in children exposed to traffic-related airborne particles. *Indoor Air* 30 (2), 235–243. [PubMed: 31743467]
- Jiang CR, Wang JL, 2010. Adjusted functional principal components analysis for longitudinal data. *Ann. Stat* 38 (2), 1194–1226.
- Jin S, Homer C, Yang L, Danielson P, Dewitz J, Li C, Zhu Z, Xian G, Howard D, 2019. Overall methodology design for the United States National Land cover database 2016 products. *Remote. Sens* 11 (24).
- Khreis H, Sanchez KA, Foster M, Burns J, Nieuwenhuijsen MJ, Jaikumar R, Ramani T, Zietsman J, 2023. Urban policy interventions to reduce traffic-related emissions and air pollution: A systematic evidence map. *Environ. Int* 172, 107805. [PubMed: 36780750]
- Knapp EA, Fink AK, Goss CH, Sewall A, Ostrenga J, Dowd C, Elbert A, Petren KM, Marshall BC, 2016. The cystic fibrosis foundation patient registry. Design and methods of a national observational disease registry. *Ann. Am. Thorac. Soc* 13 (7), 1173–1179. [PubMed: 27078236]
- Konstan MW, Morgan WJ, Butler SM, Pasta DJ, Craib ML, Silva SJ, Stokes DC, Wohl ME, Wagener JS, Regelmann WE, Johnson CA, G. Scientific Advisory, I. the and F. Coordinators of the Epidemiologic Study of Cystic, 2007. Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. *J. Pediatr* 151 (2), 134–139, 139 e131. [PubMed: 17643762]

- Louisias M, Ramadan A, Naja AS, Phipatanakul W, 2019. The effects of the environment on asthma disease activity. *Immunol. Allergy Clin. N. Am* 39 (2), 163–175.
- McArdle CE, Dowling TC, Carey K, DeVies J, Johns D, Gates AL, Stein Z, van Santen KL, Radhakrishnan L, Kite-Powell A, Soetebier K, Sacks JD, Sircar K, Hartnett KP, Mirabelli MC, 2023. Asthma-associated emergency department visits during the Canadian wildfire smoke episodes - United States, April- August 2023. *MMWR Morb. Mortal. Wkly. Rep* 72 (34), 926–932. [PubMed: 37616233]
- McGarry ME, Neuhaus JM, Nielson DW, Burchard E, Ly NP, 2017. Pulmonary function disparities exist and persist in Hispanic patients with cystic fibrosis: a longitudinal analysis. *Pediatr. Pulmonol* 52 (12), 1550–1557. [PubMed: 29082671]
- Morgan WJ, Wagener JS, Yegin A, Pasta DJ, Millar SJ, Konstan MW, i. Scientific Advisory Group and F. coordinators of the Epidemiologic Study of Cystic, 2013. Probability of treatment following acute decline in lung function in children with cystic fibrosis is related to baseline pulmonary function. *J. Pediatr* 163 (4), 1152–1157 e1152. [PubMed: 23810128]
- Nichols DP, Paynter AC, Heltshe SL, Donaldson SH, Frederick CA, Freedman SD, Gelfond D, Hoffman LR, Kelly A, Narkewicz MR, Pittman JE, Ratjen F, Rosenfeld M, Sagel SD, Schwarzenberg SJ, Singh PK, Solomon GM, Stalvey MS, Clancy JP, Kirby S, Van Dalfsen JM, Kloster MH, Rowe SM, Group PS, 2021. Clinical effectiveness of elexacaftor/tezacaftor/ivacaftor in people with cystic fibrosis: A clinical trial. *Am. J. Respir. Crit. Care Med* 205 (5), 529–539.
- Oates GR, Baker E, Collaco JM, Rowe SM, Rutland SB, Fowler CM, Harris WT, 2021. Cessation of smoke exposure improves pediatric CF outcomes: Longitudinal analysis of CF Foundation Patient Registry data. *J. Cyst. Fibros* 20 (4), 618–624. [PubMed: 34281808]
- Oates GR, Baker E, Rowe SM, Gutierrez HH, Schechter MS, Morgan W, Harris WT, 2020. Tobacco smoke exposure and socioeconomic factors are independent predictors of pulmonary decline in pediatric cystic fibrosis. *J. Cyst. Fibros* 19 (5), 783–790. [PubMed: 32081643]
- Oates GR, Schechter MS, 2016. Socioeconomic status and health outcomes: cystic fibrosis as a model. *Expert Rev. Respir. Med* 10 (9), 967–977. [PubMed: 27268142]
- Park HJ, Lee HY, Suh CH, Kim HC, Kim HC, Park YJ, Lee SW, 2021. The effect of particulate matter reduction by indoor air filter use on respiratory symptoms and lung function: a systematic review and meta-analysis. *Allergy Asthma Immunol. Res* 13 (5), 719–732. [PubMed: 34486257]
- Pfeffer PE, Mudway IS, Grigg J, 2021. Air pollution and asthma: mechanisms of harm and considerations for clinical interventions. *Chest* 159 (4), 1346–1355. [PubMed: 33461908]
- Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J, ERS Global Lung Function Initiative, 2012. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur. Respir. J* 40 (6), 1324–1343. [PubMed: 22743675]
- Schechter M, Schmidt JH, Williams R, Norton R, Taylor D, Molzhon A, 2018. Impact of a program ensuring consistent response to acute drops in lung function in children with cystic fibrosis. *J. Cyst. Fibros* 17 (6), 769–778. [PubMed: 30017327]
- Schechter MS, 2008. Patient registry analyses: seize the data, but caveat lector. *J. Pediatr* 153 (6), 733–735. [PubMed: 19014811]
- Schechter MS, 2011. Nongenetic influences on cystic fibrosis outcomes. *Curr. Opin. Pulm. Med* 17 (6), 448–454. [PubMed: 21897254]
- Schechter MS, Margolis PA, 1998. Relationship between socioeconomic status and disease severity in cystic fibrosis. *J. Pediatr* 132 (2), 260–264. [PubMed: 9506638]
- Schechter MS, Shelton BJ, Margolis PA, Fitzsimmons SC, 2001. The association of socioeconomic status with outcomes in cystic fibrosis patients in the United States. *Am. J. Respir. Crit. Care Med* 163 (6), 1331–1337. [PubMed: 11371397]
- Schluchter MD, Konstan MW, Drumm ML, Yankaskas JR, Knowles MR, 2006. Classifying severity of cystic fibrosis lung disease using longitudinal pulmonary function data. *Am. J. Respir. Crit. Care Med* 174 (7), 780–786. [PubMed: 16858011]
- Szczesniak R, Andrinopoulou ER, Su W, Afonso PM, Burgel PR, Cromwell E, Gecili E, Ghulam E, Goss CH, Mayer-Hamblett N, Keogh RH, Liou TG, Marshall B, Morgan WJ, Ostrenga JS, Pasta DJ, Stanojevic S, Wainwright C, Zhou GC, Fernandez G, Fink AK, Schechter MS, 2023. Lung

function decline in cystic fibrosis: impact of data availability and modeling strategies on clinical interpretations. *Ann. Am. Thorac. Soc* 20 (7), 958–968. [PubMed: 36884219]

Szczesniak R, Rice JL, Brokamp C, Ryan P, Pestian T, Ni Y, Andrinopoulou ER, Keogh RH, Gecili E, Huang R, Clancy JP, Collaco JM, 2020. Influences of environmental exposures on individuals living with cystic fibrosis. *Expert Rev. Respir. Med* 14 (7), 737–748. [PubMed: 32264725]

Szczesniak RD, Li D, Su W, Brokamp C, Pestian J, Seid M, Clancy JP, 2017. Phenotypes of rapid cystic fibrosis lung disease progression during adolescence and young adulthood. *Am. J. Respir. Crit. Care Med* 196 (4), 471–478. [PubMed: 28410569]

Szczesniak RD, McPhail GL, Duan LL, Macaluso M, Amin RS, Clancy JP, 2013. A semiparametric approach to estimate rapid lung function decline in cystic fibrosis. *Ann. Epidemiol* 23 (12), 771–777. [PubMed: 24103586]

United States Department of Commerce (2018). TIGER/Line Shapefiles. Retrieved 15 December 2022 from www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.

United States Environmental Protection Agency, (2015-2019). “EJScreen: environmental justice screening and mapping tool.” Retrieved December 2020, 2020, from www.epa.gov/ejscreen.

Vandenbranden SL, McMullen A, Schechter MS, Pasta DJ, Michaelis RL, Konstan MW, Wagener JS, Morgan WJ, McColley SA, Investigators and F Coordinators of the Epidemiologic Study of Cystic, 2012. Lung function decline from adolescence to young adulthood in cystic fibrosis. *Pediatr. Pulmonol* 47 (2), 135–143. [PubMed: 22241571]

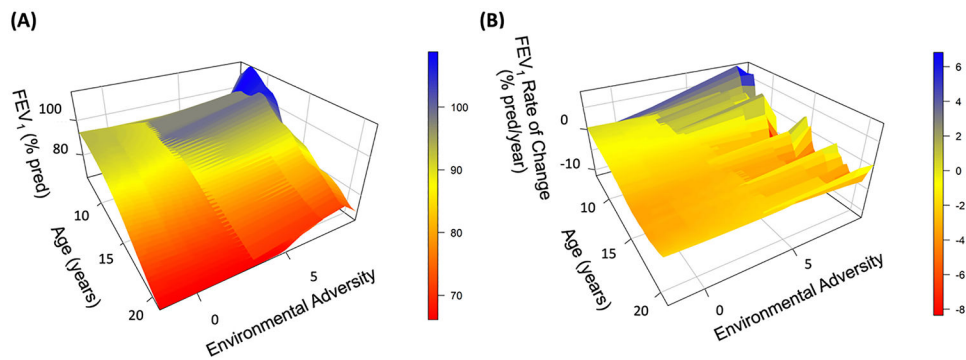


Fig. 1. Lung function trajectories by age and social-environmental adversity.

Three-dimensional plotting is used to examine how (A) lung function trajectory (FEV₁ % predicted) and (B) rate of change in lung function trajectory (% predicted/year) vary over age (the time variable, measured in years) and extent of social-environmental adversity (higher values imply more negative environmental exposure). Blue areas imply higher lung function or rate of change, while red areas indicate lower lung function or more rapid decline. Results were obtained as fitted curves from covariate-adjusted functional principal components analysis (see Study Design and Methods section). FEV₁ (% pred) indicates forced expiratory volume in 1 s of % predicted. Sub-plots by age are shown as supplemental material (e-Figs. 4 and 5).

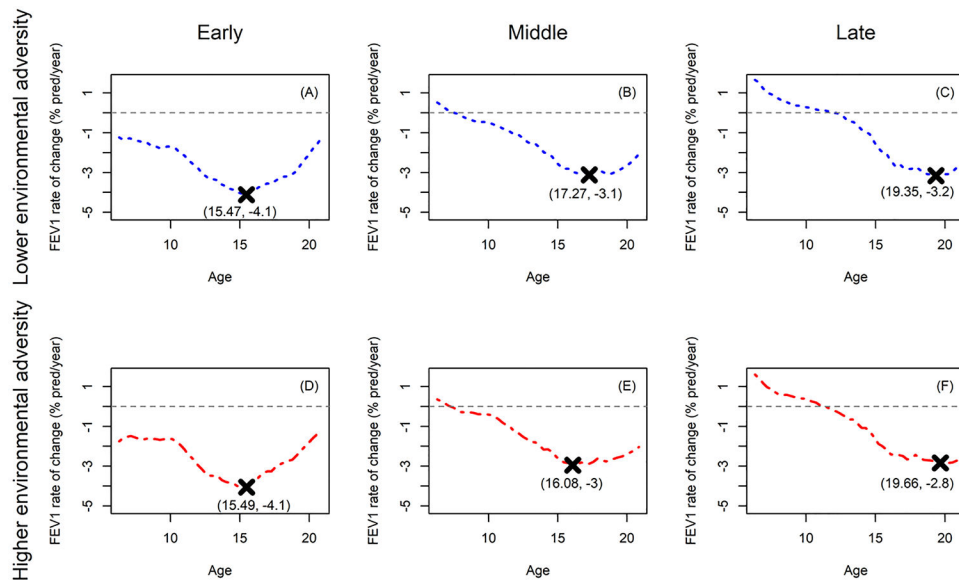


Fig. 2. Peak decline in lung function over age according to social-environmental phenotype. Average rates of change (y-axis) over age (the time variable, measured in years) are shown for early (A and D), middle (B and E), and late (C and F) environmental phenotypes of rapid decline. Rate of change is shown for each phenotype sub-grouped by degree of social-environmental adversity (lower and higher correspond to blue dashed and red dot-dashed curves, respectively). “X” is used to mark coordinates for the average timing of peak decline and estimated age at which it occurred. Results were obtained by differentiating the fitted curves from covariate-adjusted functional principal components analysis (see Study Design and Methods). FEV₁ (% pred/year) indicates annualized rate of change forced expiratory volume in 1 s of % predicted.

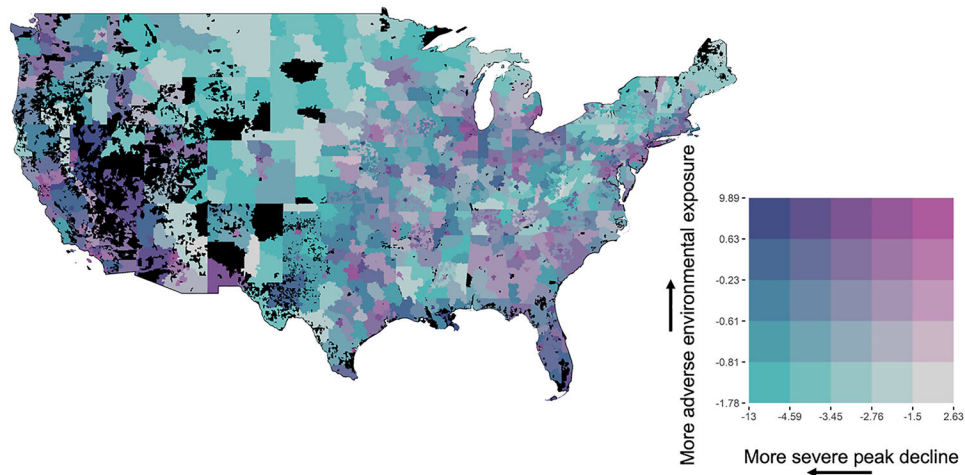


Fig. 3. Geographic variation of social-environmental adversity and rapid lung function decline. Bivariate quantities are shown representing extent of social-environmental adversity and severity of peak decline in lung function in the continental U.S. Higher values of social-environmental adversity imply more negative social and/or environmental exposure levels (vertical arrow pointing upward), and more negative values of severity of peak decline imply greater maximal loss of lung function (horizontal arrow pointing leftward). Dark blue (upper left corner of the grid) represents extremely rapid decline and worst environmental adversity, while gray (lower right corner of the grid) corresponds to at or below average rate of decline (-1.5% predicted/year) with least degree of social-environmental adversity. Results were obtained from two-stage cluster analysis (see Study Design and Methods). Estimates as shown on the map are aggregated to three-digit ZIP Codes and displayed using HIPAA Safe Harbor Guidelines. A three-digit ZIP Code was colored black if it did not contain either (1) a population of at least 20,000 residents or (2) a sufficient number of residents in the analysis cohort to make an estimate.

Table 1

Demographic and clinical characteristics, overall and across social-environmental phenotypes.

	Overall cohort (n = 24,228)	Early (n = 6,057)	Middle (n = 12,114)	Late (n = 6,057)
F508del mutation*				
Homozygous	11,618 (48.0 %)	2,887 (47.7 %)	6,006 (49.6 %)	2,725 (45.0 %)
Heterozygous	8,711 (36.0 %)	1,983 (32.7 %)	4,363 (36.0 %)	2,365 (39.0 %)
Neither/unknown	3,899 (16.0 %)	1,187 (19.6 %)	1,745 (14.4 %)	967 (16.0 %)
Male*	12,367 (51.0 %)	2,769 (45.7 %)	6,325 (52.2 %)	3,273 (54.0 %)
Race ^b				
White	22,486 (92.8 %)	5,589 (93.4 %)	11,241 (92.8 %)	5,656 (93.4 %)
Non-White	1,742 (7.2 %)	468 (7.7 %)	873 (7.2 %)	401 (6.6 %)
Hispanic ethnicity*	1,973 (8.1 %)	650 (10.7 %)	916 (7.6 %)	407 (6.7 %)
Diagnosis age ^{b,c}	2.0 (3.7)	1.55 (2.99)	2.00 (3.71)	2.43 (4.05)
Baseline age ^{a,b,c}	8.6 (6.0 - 20.4)	9.2 (3.8)	8.5 (3.6)	8.3 (3.4)
Birth year ^{a,b,c}	1990 (1980–2010)	1990 (1980–2010)	2000 (1980–2010)	2000 (1980–2010)
Baseline FEV ₁ (% predicted) ^{a,b,c}	89.2 (21.6)	66.1 (19.6)	89.9 (15.4)	107 (13.9)
Non-private or no insurance				
At baseline*	12,160 (50.2 %)	3,625 (59.8 %)	6,000 (49.5 %)	2,535 (41.9 %)
Ever during follow-up*	19,272 (79.5 %)	5,064 (83.6 %)	9,591 (79.2 %)	4,617 (76.2 %)
Microbiology				
Pa				
At baseline*	4,495 (18.6 %)	1,612 (26.6 %)	2,153 (17.8 %)	730 (12.1 %)
Ever during follow-up*	19,449 (80.3 %)	5,564 (91.9 %)	9,556 (78.9 %)	4,329 (71.5 %)
MRSA				
At baseline ^a	1,247 (5.1 %)	366 (6.0 %)	639 (5.3 %)	242 (4.0 %)
Ever during follow-up ^a	10,897 (45.0 %)	2,997 (49.5 %)	5,506 (45.5 %)	2,394 (39.5 %)
CFRD status				
Impaired				
Glucose				
Tolerance				
At baseline	67 (0.3 %)	22 (0.4 %)	33 (0.3 %)	12 (0.2 %)
Ever during follow-up*	3,220 (13.3 %)	736 (12.2 %)	1,596 (13.2 %)	888 (14.7 %)
CFRD diagnosis				
At baseline ^{a,b}	117 (0.5 %)	44 (0.7 %)	48 (0.4 %)	25 (0.4 %)
Ever during follow-up*	5,880 (24.3 %)	2,347 (38.7 %)	2,577 (21.3 %)	956 (15.8 %)
Pancreatic Enzymes	22,955 (94.7 %)	5,987 (98.8 %)	11,471 (94.7 %)	5,497 (90.8 %)
Ivacaftor [#]				
Ever during follow-up*	1,067 (4.4 %)	107 (1.8 %)	537 (4.4 %)	423 (7.0 %)

	Overall cohort (n = 24,228)	Early (n = 6,057)	Middle (n = 12,114)	Late (n = 6,057)
Lumacaftor/Ivacaftor [#]				
Ever during follow-up [*]	3,930 (16.2 %)	697 (11.5 %)	2,214 (18.3 %)	1,019 (16.8 %)
PEX frequency in year prior to baseline				
None <i>a.b.c</i>	21,130 (87.2 %)	5,030 (83.0 %)	10,551 (87.1 %)	5,549 (91.6 %)
1 <i>a.b.c</i>	1,049 (4.3 %)	327 (5.4 %)	555 (4.6 %)	167 (2.8 %)
2 <i>a.b.c</i>	287 (1.2 %)	84 (1.4 %)	148 (1.2 %)	55 (0.9 %)
3 or more <i>a.b.c</i>	1,762 (7.3 %)	616 (10.2 %)	860 (7.1 %)	286 (4.7 %)
Hospital visits in the year prior to baseline visit				
None <i>a.b.c</i>	20,115 (83.0 %)	4,841 (79.9 %)	10,000 (82.5 %)	5,274 (87.1 %)
1 <i>b.c</i>	1,147 (4.7 %)	305 (5.0 %)	605 (5.0 %)	237 (3.9 %)
2 <i>b</i>	463 (1.9 %)	130 (2.1 %)	237 (2.0 %)	96 (1.6 %)
3 or more <i>a.b.c</i>	2,503 (10.3 %)	781 (12.9 %)	1,272 (10.5 %)	450 (7.4 %)
Smoke exposure ^{##}				
At baseline [*]	1,296 (5.3 %)	266 (4.4 %)	292 (4.8 %)	738 (6.1 %)
Ever during follow-up [*]	8,130 (33.6 %)	1,845 (30.5 %)	4,227 (34.9 %)	2,058 (34.0 %)
Lung transplant during follow-up [*]	1810 (7.5 %)	1,251 (20.7 %)	523 (4.3 %)	36 (0.6 %)
Alive through follow-up [*]	20,380 (84.1 %)	3,519 (58.1 %)	10,982 (90.7 %)	5,879 (97.1 %)

Mean (SD) and n (%) are reported for continuous and categorical variables, respectively. P-values from Wilcoxon rank sum test or chi-square test. Statistical significance of comparisons (P-value < 0.05) marked as

^aEarly versus Middle;

^bEarly versus Late;

^cMiddle versus Late for continuous variables and

^{*}evidence of overall association between phenotype and categorical variable.

[^]Indicates insufficient sample size or censoring prohibiting standard statistical comparison. Abbreviations include CFRD = cystic fibrosis-related diabetes; MRSA = methicillin-resistant *Staphylococcus aureus*; *Pa* = *Pseudomonas aeruginosa*; PEX = pulmonary exacerbation.

[#]Baseline use of modulator therapies was suppressed for patient privacy purposes due to low cell counts (< 5 subjects).

^{##}Types of smoke exposure (firsthand, secondhand and within household) were combined due to low cell counts.

Table 2

Baseline geomarker characteristics, overall and across social-environmental phenotypes.

	Overall cohort (N = 24,228)	Early (n = 6,057)	Middle (n = 12,114)	Late (n = 6,057)
Deprivation index ^{a,b,c}	0.341 (0.0983)	0.359 (0.100)	0.340 (0.0972)	0.324 (0.0950)
Total crime ^{a,b,c}	87.3 (59.7)	90.4 (57.8)	86.6 (59.2)	85.6 (62.5)
Landcover				
% Green ^{b,c}	83.4 (20.9)	83.1 (21.5)	83.6 (20.8)	83.2 (20.3)
% Impervious ^{b,c}	15.9 (18.1)	16.1 (18.6)	15.7 (18.1)	16.2 (17.8)
% Tree canopy	26.8 (21.6)	27.1 (22.3)	26.7 (21.4)	26.6 (21.4)
Respiratory hazard index ^a	1.57 (0.811)	1.59 (0.791)	1.56 (0.815)	1.58 (0.823)
Ozone concentration ^{a,b}	46.2 (6.69)	46.4 (6.85)	46.1 (6.62)	45.9 (6.65)
PM _{2.5} Concentration	9.64 (1.58)	9.68 (1.57)	9.63 (1.58)	9.63 (1.57)
Traffic proximity ^c	74.4 (128)	75.3 (121)	73.1 (131)	76.1 (128)
Diesel particulate matter ^c	0.695 (0.615)	0.696 (0.606)	0.689 (0.623)	0.705 (0.608)
Primary roadways				
Length	19100 (27600)	19000 (27800)	19400 (28000)	18400 (26300)
Density	1.24 (2.75)	1.30 (2.87)	1.21 (2.70)	1.23 (2.72)
Secondary Roadways				
Length ^{b,c}	59300 (73200)	63100 (76700)	59900 (74600)	54400 (66400)
Density ^{a,b}	2.54 (4.98)	2.61 (4.45)	2.55 (5.41)	2.46 (4.56)

Mean (SD) are reported. P-values from Wilcoxon rank sum test. Statistical significance of comparisons (P-value < 0.05) marked as

^aEarly versus Middle;^bEarly versus Late^cMiddle versus Late.

Table 3

Clinical translation of social-environmental phenotypes.

Phenotype	Low social-environmental adversity ^a	High social-environmental adversity ^b
Early	<ul style="list-style-type: none"> Proactive indication of clinical risk factors identified in this study, including older age, established lung infections, comorbidities, history of PEx; Hispanic ethnicity (McGarry, Neuhaus et al. 2017); possibly ultra-rare mutations (those with no F508del alleles, modulator ineligible) 	<ul style="list-style-type: none"> Prone to higher community deprivation and total crime, may consider more frequent psychosocial assessment (Oates and Schechter 2016) Also close proximity to primary roadways, may consider HEPA filters in the home. Needs further study; asthma findings have been mixed (Park, Lee et al. 2021)
Middle	<ul style="list-style-type: none"> Routine monitoring of microbiology and PFT per care guidelines (Foundation 2022) 	<ul style="list-style-type: none"> Most susceptible to respiratory hazards, may consider proper masking during spikes in air pollution levels to avoid triggering PEx (e.g.: in CF (Goeminne, Kicinski et al. 2013) or asthma severity during wildfires (McArdle, Dowling et al. 2023))
Late	<ul style="list-style-type: none"> Avoid untreated declines in individuals with higher lung function (Morgan, Wagener et al. 2013) 	<ul style="list-style-type: none"> High levels of traffic-related air pollution, may require air quality policies in urban areas (Khreis, Sanchez et al. 2023)

^aIncludes measures to improve lung function that have been identified in the absence of social and environmental exposure data.

^bPotential clinical actions are based on results from this study and other literature; interventional studies are needed for cystic fibrosis. Purple arrow denotes that these interventions could be applied among all three decline phenotypes depending on personalized factors. Each horizontal arrow and light-to-dark color shading illustrates, for a given phenotype, that the interventions specific to the category of low social-environmental adversity could be layered for deployment with interventions in the corresponding high social-environmental adversity level. Abbreviations: PEx = pulmonary exacerbation; PFT = pulmonary function testing.