TEMPERATURE AND LUNG FUNCTION IN CYSTIC FIBROSIS

by
Joseph Michael Collaco

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

March 2016

ABSTRACT

Cystic fibrosis (CF) is the most common life-limiting genetic disorder in the United States, and is characterized by progressive obstructive lung disease. The severity of lung disease can be highly variable even among individuals with identical mutations in the disease-causing gene, *CFTR*. Half of the variation in CF lung function is attributable to environmental/stochastic factors, thus identifying environmental factors that affect lung function is crucial in ameliorating the disease burden.

Using data from the CF Twin-Sibling Study and two national registries, we previously demonstrated that warmer annual ambient temperatures were associated with lower lung function after adjusting for *CFTR* genotype and demographic/environmental factors. However, the mechanisms behind this association remain to be elucidated. Thus, the objectives of this thesis were to identify mechanisms by which ambient temperature affects lung function, specifically through assessments of common CF respiratory pathogens using databases of subjects with CF (Chapter 2), tests for gene-environment interactions using genomic data from the CF Twin-Sibling Study (Chapter 3), and assessment of whether this phenomenon exists in the general population using NHANES databases (Chapter 4).

Taken together, the results from Chapters 2 and 4 of this study potentially explain 40-83% of the association between lower lung function and warmer temperatures in CF and identify some mechanisms that underlie this association. The mediation analyses in Chapter 2 in two separate populations of individuals with CF ascribed 12-43% of the association in CF to 3 respiratory pathogens. Chapter 4 also confirmed the presence of this association within the general population with effect sizes of 0.7% and 1.0%

predicted FEV_1 decrease for every 10°F increase in mean annual ambient temperature in the two NHANES general population cohorts and suggested that 28-40% of the association seen in CF can be ascribed to a mechanism also present in the general population. Chapter 3 did not identify any specific gene-environment interactions with genetic modifiers of CF lung disease or genetic modifiers of lung function in the general population. Our results have important implications for both individuals with CF with respect to their healthcare and the general population with respect to population health and climate change.

Thesis Advisor:	Lawrence J. Appel, M.D., M.P.H.
Mentor:	Garry R. Cutting, M.D.
Committee Members:	Terri H. Beaty, Ph.D.
	David W. Dowdy, M.D., Ph.D., Sc.M.
Readers:	Priya Duggal, Ph.D., M.P.H.
	John McGready, Ph.D., M.S.
	Ernst W. Spannhake, Ph.D.

PREFACE

There are a number of people without whom this dissertation would not have been written, and to whom I am greatly indebted. For their guidance, time, and patience, I wish to thank my thesis advisor, Dr. Larry Appel, and my research mentor of over a decade, Dr. Garry Cutting, for their sage advice, my thesis committee members, Drs. Terri Beaty and David Dowdy, for willingly taking on the task of serving as readers, Drs. John McGready, Priya Duggal, and William Spannhake, for assistance with data processing and analysis, Dr. Scott Blackman, Dr. Deanna Green, Karen Raraigh, and Kathleen Naughton, for invaluable assistance with the logistics of scheduling and meeting deadlines, Fran Burman and Trish Cornwall, and for additional mentorship and protecting my time to accomplish this, Drs. Sharon McGrath-Morrow and Pam Zeitlin. This work also would not have been possible without funding from the Cystic Fibrosis Foundation and Gilead Sciences, Inc, and therapeutic support from Ole Kirk Christiansen.

I dedicate this thesis to my family, including my very supportive in-laws, the sisthren, Anne, Lisa, and Renu, my wife, Aarti, for being there through the long working hours, and lastly my parents for their unconditional love and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Section 1.1. Cystic Fibrosis

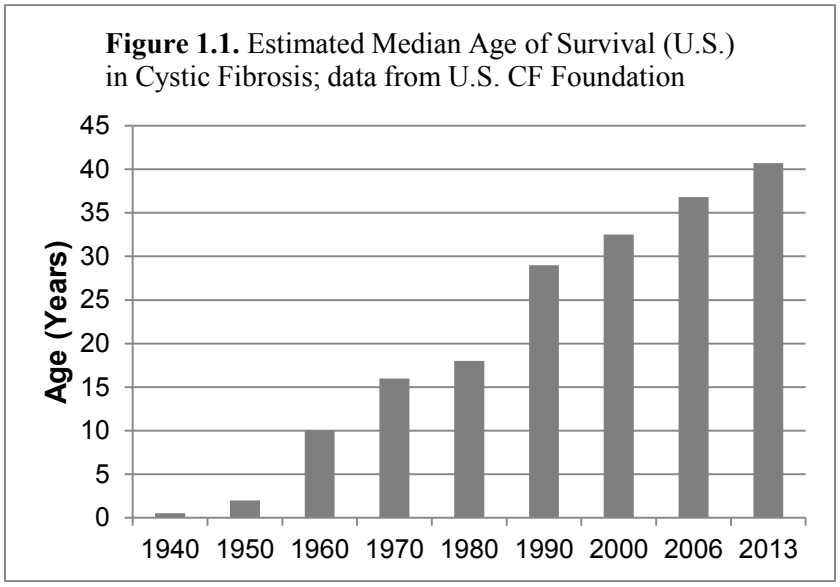


Figure 1.1. Estimated Median Age of Survival (U.S.) in Cystic Fibrosis; data from U.S. CF Foundation

Cystic fibrosis is one of the most common life-limiting genetic disorders affecting over 30,000 people in the United States and over 70,000 worldwide (www.cff.org). Although therapeutic advances have dramatically increased the median age of survival (Figure 1.1), life expectancy is still approximately half that of the general population. The underlying genetic defect in cystic fibrosis was identified in 1989 and consists of mutations in the *CFTR* (Cystic Fibrosis Transmembrane conductance Regulator) gene on chromosome 7, an ABC transporter class chloride ion channel found in epithelial cells.[1] Over 2000 mutations in this gene have been reported (http://www.genet.sickkids.on.ca), however not all of these mutations are disease causing (e.g., p.Met470Val). Specific mutations are correlated with certain aspects of disease severity, such as pancreatic insufficiency, but not others, such as lung function.[2-4] Inheritance is mediated in an autosomal recessive manner; carriers are generally asymptomatic. Given the prevalence of carriers in the population (1:30 U.S. Caucasians), mutations are suspected to have a heterozygote advantage, but the specific evolutionary pressure remains controversial.[5-7]

Clinically, cystic fibrosis manifests in multiple epithelial organs (Figure 1.2) as the mutated channel protein is produced in reduced numbers, degraded more rapidly, or functions poorly, leading to thickened secretions. Blockage of pancreatic ducts and subsequent injury leads to exocrine and endocrine pancreatic disease, resulting malabsorption, malnutrition, and diabetes. Thick respiratory secretions that are poorly cleared from the lung and sinuses lead to colonization by respiratory pathogens resulting in acute episodes of bronchitis, pneumonia, and sinusitis, as well as chronic, progressive lung disease. It is the latter decline in lung function that currently results in 90% of the mortality seen in this disorder. The milieu of respiratory organisms colonizing the airway does change with age (Figure 1.3), with certain organisms, such as *Pseudomonas aeruginosa*,[8,9] *Burkholderia cepacia* complex,[10-12] and methicillin-resistant *Staphylococcus aureus* (MRSA)[13-15] being associated with worse prognosis.

Figure 1.2. Common Manifestations of Cystic Fibrosis

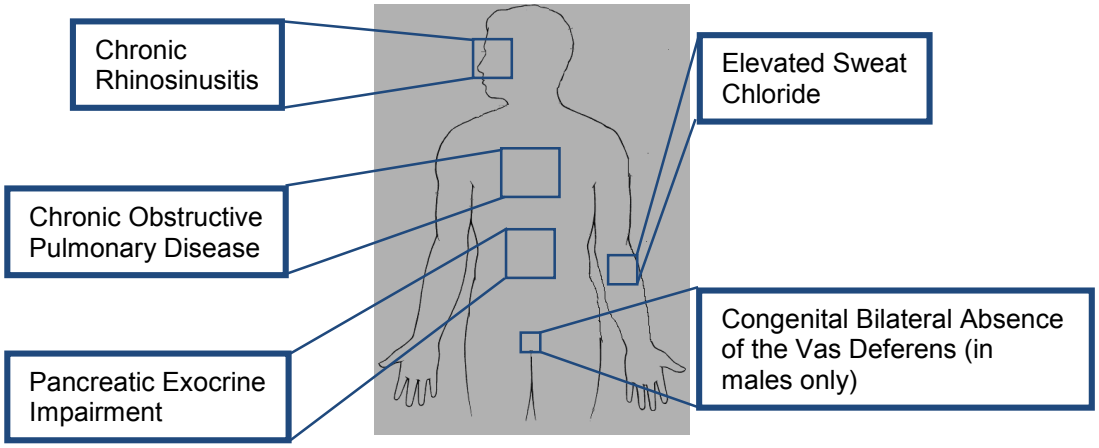
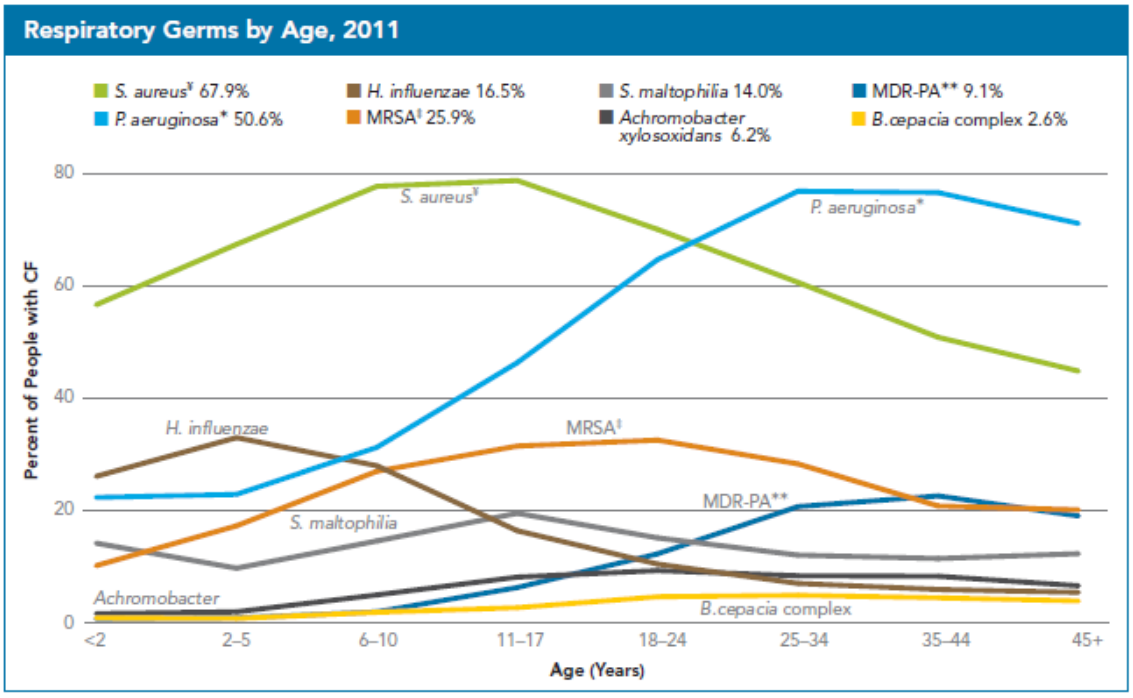


Figure 1.3. Common Respiratory Pathogens in CF by Age (From the U.S. CF Foundation Patient Registry 2011 Report)



*P. aeruginosa *includes people with MDR-PA.*
***MDR-PA is multi-drug resistant* Pseudomonas aeruginosa (P. aeruginosa).**
¥S. aureus *includes people with MRSA.*
‡*MRSA is methicillin-resistant* Staphylococcus aureus (S. aureus).

As is seen in Figure 1.3, given enough time, most people with CF are colonized with *P. aeruginosa.* Current standard of care is to attempt eradication at the first positive culture for *P. aeruginosa*. Once *P. aeruginosa* colonizes the respiratory tract, it frequently mutates into a "mucoid" strain, which confers the organism certain protection to clearance.[16,17] Other disease manifestations of cystic fibrosis include meconium ileus, nasal polyps, cirrhotic liver disease, and male infertility owing to absence of the vas deferens.

Optimal care for people with CF in the U.S. currently includes a minimum of 4 visits annually at an accredited or affiliate CF center with a multidisciplinary team, which

may consist of a physician (frequently a pulmonologist), nutritionist, respiratory therapist, social worker, CF nurse or case manager, physical therapist, pharmacist, and psychologist. At these routine visits, spirometry is obtained as a measure of lung function and respiratory cultures (expectorated sputum or throat swab) are obtained for qualitative surveillance. These data can serve as longitudinal measures for clinical research studies. With regards to lung function, it is more likely that spirometry measures underestimate true lung function owing to inadequate patient effort, poor technique, or illness, rather than overestimate. With regards to respiratory cultures, throat swabs, in particular, may not accurately represent lower respiratory flora.[18,19]

In terms of maintenance therapies, most people with CF have exocrine pancreatic insufficiency from birth and require pancreatic enzymatic supplements with every meal, snack, and/or supplemental tube feed. Most people with CF also require a complex regimen of airway clearance consisting of vibrating chest vests, aerosolized medications, etc. to aid with clearing the lower airways of thick secretions. They may also be receiving prophylactic antibiotics to decrease bacterial load in the airways. Estimates suggest that completing all recommended therapies may require at least 1-2 hours daily, a significant burden for patients and their families.[20]

People with CF often experience respiratory exacerbations that manifest with fevers, increased cough, sputum production, weight loss, wheezing, increased fatigue, and/or an abrupt decline in lung function. Typical management includes increased frequency of airway clearance and either oral or intravenous antibiotics. Although patients do regain some lung function after an exacerbation, they frequently do not return to their pre-exacerbation baseline.[21]

Lastly, care is continuing to improve with innovative new therapies.[22,23] In January 2012, the U.S. Food and Drug Administration approved the first medication, ivacaftor, to target the defect of CF. Although the FDA-approved indications for use of the CFTR potentiator ivacaftor includes ten *CFTR* mutations,[24] these mutations account for a small fraction of the CF population.[25] The use of drugs such as ivacaftor and other *CFTR* potentiators and correctors is expanding. For example, the combination of ivacaftor and the corrector lumacaftor for patients homozygous for the most common *CFTR* mutation, *F508del*, was just approved in July 2015.[26]

Section 1.2. Research Significance

Cystic fibrosis (CF) is the most common life-limiting genetic disorder in the United States, affecting over 30,000 individuals. Despite advances in nutritional and respiratory care, the median age of survival in 2013 (40.7 years) was approximately half that of the general population. Clinically, CF is characterized by chronic progressive obstructive lung disease, which is responsible for the majority of morbidity and mortality associated with this disorder. Progressive lung disease is accelerated by recurrent respiratory infections with irrevocable loss of lung function despite effective therapy for symptoms.[21,27] Although the causal gene, the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, has been known for over 20 years,[1] it has become evident that significant variation in the severity CF lung disease exists even among individuals with identical mutations.[2-4] Twin and sibling studies of lung function in individuals with CF indicate that approximately half of the variation in lung function is attributable to genetic factors, but half is due to environmental and/or stochastic (random) factors.[28-30]

As survival is correlated with severity of lung disease,[31,32] understanding the environmental factors that affect lung function is crucial in ameliorating the disease and financial burden imposed by CF. Although a number of studies have demonstrated associations between lung function and individual environmental factors, including secondhand smoke exposure,[33-37] air pollution,[38] household income,[31-33] maternal education,[39] and insurance status,[39-42] limited analyses have been performed to determine the relative contribution of these factors.[39,43] Furthermore, most studies have focused on environmental factors operating at a household level. With the exception of air pollution,[38] there has been limited work examining environmental factors operating on a geographic scale, such as climate.

Using data from the U.S. CF Twin-Sibling Study, we previously demonstrated that warmer annual ambient temperatures were associated with lower lung function; we subsequently independently replicated this finding in two national registries in the United States and Australia.[44] We also tested the contribution of temperature relative to other environmental factors, and found that temperature remained significantly associated with lung function even after adjusting for *CFTR* genotype, demographic factors, socio-economic factors, and other geographic factors, such as relative humidity. The mechanisms by which temperature affects lung function in individuals with CF remain to be elucidated.

Ambient temperature has been demonstrated to be associated with the prevalence of both infectious and non-infectious diseases,[45] particularly with vector-borne infectious diseases; however, our replicated association between ambient temperature and lung function is novel. Studying this association may yield new insights into the mechanisms

by which geographic factors, such as climate, affect human health, and may have ramifications for informing healthcare policy. Thus, this thesis may inform our knowledge of CF respiratory physiology and may increase our understanding of the epidemiology of common CF respiratory pathogens (Chapter 2), may identify gene-environment interactions which may contribute to the development of CF lung disease (Chapter 3), and may improve our understanding of respiratory pathophysiology in the general population (Chapter 4).

More specifically in Chapter 2, the expected findings will allow respiratory and infectious clinical outcomes from study sites and CF centers across a wide geographic area to be more accurately compared without temperature bias. The clinical care of patients with CF has benefited from one of the first disease-specific data registries, the U.S. Cystic Fibrosis Foundation Patient Registry. Benchmarking using aggregate data by individual CF care centers by the U.S. CF Foundation Patient Registry has spurred on widespread efforts for quality improvement. Based on our preliminary data, we would estimate a hypothetical 18 year old male with CF (Height: 175cm; Non-Hispanic White) with an NHANES[46] FEV_1 of 73.5% percent living in a climate with colder temperatures would be expected to have an FEV_1 of 66.1% had he resided in a 30 degree (°F) warmer climate, which would be a clinically relevant difference in lung function. Given that annual ambient temperature is not within patients' and clinical providers' control, geography may need to be considered when patients, clinicians, and payers compare the performance of different CF Care Centers. Furthermore, ambient temperature may also need to be considered in the design of multi-center CF studies of lung function as well as clinical trials that encompass broad geographic areas. Because of these considerations

and the ongoing need to improve outcomes, it is important to determine by which means temperature affects lung function as this may lead to improved means of risk stratification and determining clinical trajectories. It is possible that information from this study may guide regionally based initiatives or management strategies to improve outcomes as well.

This study represents one of the first efforts to identify gene-environment interactions with a replicated environmental modifier of lung function (i.e., temperature) that operates on a geographic scale. The proposed work may serve to inform other similar efforts. More specifically, in Chapter 3, identifying modifier gene-temperature interactions would provide pilot data to explore why such significant variation in CF lung disease exists among individuals with identical *CFTR* mutations, and will allow future tests of genetic modifiers to be more accurately interpreted in light of a patient's geographic location.

Lastly in Chapter 4, the knowledge of whether temperature is associated with lung function in the general population may inform our understanding of pathophysiology of lung disease not only for CF patients, but also for lung function in the general population and among other individuals with obstructive lung diseases, such as asthma and COPD.

Section 1.3. Approach for Chapter 2

To identify (a) which common respiratory pathogens in CF are associated with temperature and (b) the extent to which these organisms explain the relationship between temperature and lung function.

Introduction: Determining factors that influence the acquisition and the timing of acquisition of pathogens is essential step in improving outcomes for individuals with CF. Owing in part to impaired airway clearance, acute infections and colonization with respiratory pathogens are common in CF. Specific respiratory pathogens that have been demonstrated to result in lower lung function or increased mortality in CF include *Pseudomonas aeruginosa*,[8,9] *Burkholderia cepacia* complex,[10-12] and methicillin-resistant *Staphylococcus aureus* (MRSA).[13-15] Furthermore, earlier acquisition of, at least, *P. aeruginosa* as detected on respiratory cultures is associated with the development of more severe lung disease.[47]

The objective of Chapter 2 is to determine which respiratory pathogens in CF are associated with temperature and if associations between temperature and infection exist, whether they explain the association between temperature and lung function. Assessing the association between temperature and infection will be accomplished through multivariate regression. The contribution of any such association to lung function will be assessed through mediation analysis.[48,49] The primary study population will be derived from the U.S. Twin-Sibling Study with replication in the CF Foundation Patient Registry.

Preliminary Data: Our preliminary data demonstrated an association between warmer temperatures and lower lung function in three independent samples of individuals with CF (Table 1.1). Samples were derived from the U.S. CF Twin-Sibling Study, the U.S. CF Foundation Patient Registry, and the Australian CF Data Registry. Associations between temperature and lung function were tested for using stepwise linear regression clustered for family (Twin-Sibling Study) or CF care center (U.S. Registry). The adjusted regression coefficients indicated that lung function is ~3 percentile points lower for every

10°F increase in annual ambient temperature. As an example, there is a 30°F difference in mean annual temperature between Florida and Maine. This association was independent of socio-economic factors such as insurance status or estimated household income as well as *CFTR* genotype as measured by number of *F508del* alleles. The association seen in the Twin-Sibling Study sample was replicated in the U.S. Registry sample. The association trended towards significance in the Australian Registry sample and the 95% C.I. did overlap the estimates from the U.S. populations. Differences between the analyses in the United States versus Australia may be secondary to a reduced absolute range of temperatures in Australia or a shift in Australia to warmer temperatures.

Table 1.1. Preliminary Data: Higher Ambient Temperatures are Associated with Lower Lung Function (CF-specific FEV_1)*

Study	**U.S. Twin-Sibling Study**	**U.S. CF Registry**	**Australian CF Registry**
N	1313	15174	1791
Adjusted Co-efficient for Temperature (per 10°F) [95%CI]	-3.4 [-5.7, -1.0]	-3.1 [-4.1, -2.1]	-2.3 [-4.7, 0.1]
Temperature Co-efficient *p* Value	**0.005**	**<0.001**	0.057

*Regressions were adjusted for insurance status and age at the time of pulmonary function testing in the U.S. samples, and age for the Australian sample.

Preliminary data also demonstrated an association between a higher prevalence of a significant CF respiratory pathogen, *Pseudomonas aeruginosa*, and warmer ambient temperature (Table 1.2). Independent samples of subjects with CF included the previous three samples as well as a prospectively recruited sample from Australia and New Zealand through the Australian CF BAL (ACFBAL) study. These associations persisted even after adjustment for *CFTR* genotype, age at time of last respiratory culture, and age at diagnosis. In addition, an earlier age of acquisition of *P. aeruginosa* was observed with warmer temperatures in the two populations where longitudinal data were present (Twin-

Sibling Study, ACFBAL). The mean ages of acquisition were younger in the ACFBAL study compared to Twin-Sibling Study likely owing to ascertainment method differences. Specifically, the ACFBAL study included a complete record of respiratory cultures from early infancy as well as the more frequent use of bronchoscopy to obtain lower airway culture samples, which may be more sensitive than the typical throat swabs used in routine clinical collection as in the Twin-Sibling Study. It is also possible that the younger ages of acquisition seen in the ACFBAL study may reflect just an earlier age of acquisition unique to Australasia. No association with temperature was seen with mucoid *P. aeruginosa* (Twin-Sibling Study) or *Burkholderia cepacia* complex (Twin-Sibling Study, U.S. and Australian Registries).

Table 1.2. Preliminary Data: Higher Annual Ambient Temperatures are Associated with a Higher Prevalence and Earlier Age of Acquisition of *P. aeruginosa*

			Stratification by Temperature Quartiles (°F)				
Variable	**Study Sample**	**N**	**Coldest**	**← Temperate →**		**Warmest**	**Chi square or ANOVA *p* value**
U.S. Temperature Quartiles		-	< 49.2	49.2 – 52.0	52.1 – 58.1	> 58.1	-
Australian Temperature Quartiles		-	< 59.1	59.1 – 63.0	63.1 – 65.7	> 65.7	-
P. aeruginosa (% Positive)	U.S. Twin-Sib Study	1372	85.3	85.0	88.4	92.9	**0.005**
	U.S. Data Registry	15,532	57.6	60.8	59.1	62.8	**<0.001**
	Aus. Data Registry	1791	72.8	81.2	82.1	83.3	**0.001**
	Aus. BAL Study	166	39.6	65.1	70.6	73.6	**0.002**
P. aeruginosa Age of Acquisition (Yrs)	U.S. Twin-Sib Study	664	7.6 ± 7.5	6.2 ± 6.5	6.3 ± 5.8	6.3 ± 5.0	**0.04**
	Aus. BAL Study	100	2.6 ± 1.5	2.7 ± 1.5	2.6 ± 1.4	1.8 ± 1.3	**0.04**

To examine the relationship between infection with *P. aeruginosa*, temperature, and lung function, we performed mediation analysis. As can be seen in Table 1.3, the inclusion of *P. aeruginosa* in the models of lung function resulted in a decrease in the magnitude of the coefficient for temperature in all three study populations. The relative decrease in the magnitude of the coefficient for temperature is the portion of the association between temperature and lung function that could be ascribed to *P. aeruginosa* (Twin-Sibling Study: 15%; U.S. Registry: 6%; Australian Registry: 22%). It should be noted that these results could also represent confounding by another unmeasured variable or collider-stratification bias.

Table 1.3. Preliminary Data: Regression Analyses for Lung Function: Assessing the Mediation Effect of *P. aeruginosa*[*]

	U.S. Twin-Sibling Study		**U.S. Registry**		**Australian Registry**	
Co-efficient [95%CI] (*p* value)	**Base Model**	**Model with *P. aeruginosa***	**Base Model**	**Model with *P. aeruginosa***	**Base Model**	**Model with *P. aeruginosa***
Multivariate Sample n	1313	1313	15174	15174	1791	1791
Temperature (°F)	-0.34 [-0.57, -0.10] (0.005)	-0.29 [-0.52, -0.06] (0.014)	-0.31 [-0.41, -0.21] (<0.001)	-0.29 [-0.39, -0.19] (<0.001)	-0.23 [-0.47, 0.01] (0.06)	-0.18 [-0.41, 0.06] (0.15)
P. aeruginosa (0=Negative, 1 = Positive)	-	-10.72 [-14.19, -7.24] (<0.001)	-	-6.92 [-7.95, -5.88] (<0.001)	-	-8.04 [-10.98, -5.10] (<0.001)

[*]Insurance status and age at the time of pulmonary function testing were also significant predictors of lung function for both U.S. samples, but not the Australian Registry sample.

Study populations: Although more detailed study population descriptions will be included in the relevant sections of Chapters 2, 3, and, 4, a brief summary of these populations follows (Table 1.4).

Cystic Fibrosis Twin-Sibling Study: The dataset included 1730 geocoded individuals with pulmonary function test and respiratory culture data. The study is based at Johns Hopkins University (PI: Dr. Garry Cutting) and has recruited families primarily

from the U.S. between 2000 and 2013. It is estimated that ~90% of the twins with CF and ~60% of the siblings with CF in the U.S. have been recruited. Collected data included questionnaires including environmental exposures, clinical data including respiratory cultures and pulmonary function tests, and DNA (blood) samples. Clinical and environmental data for subjects receiving care from U.S. CF care centers are supplemented on an annual basis using data from the U.S. CF Foundation Patient Registry. The last update covered through calendar year 2011.

U.S. CF Foundation Patient Registry: The U.S. Registry is a partnership between patients, CF care centers, and the CF Foundation to collect clinical and environmental data on almost all patients receiving care at CF care centers in the U.S. This patient registry represents the largest database of individuals with CF in the world. The data used in this study included a download of the relevant de-identified data for the calendar year 2007, comprising 24,799 individuals. Subjects recruited in the Twin-Sibling Study population were excluded from analyses using the U.S. Registry population. The dataset included 15,174 geocoded individuals with pulmonary function test and respiratory culture data.

Table 1.4. Study Sample Details

		CF Twin-Sibling Study	U.S. CF Patient Registry	NHANES
Samples to be used	Chapter 2 (Respiratory Pathogens)	√	√	
	Chapter 3 (Gene-Environment Interaction)	√		
	Chapter 4 (General Population)			√
Non-Clinical Data	Demographic and Climate Data	√	√	√
	Genetic Data	√		
Clinical Data	Cross-Sectional Lung Function Data	√	√	√
	Longitudinal Lung Function Data	√	Not available	Not available
	Cross-Sectional Respiratory Culture Data	√	√	Not applicable
	Longitudinal Respiratory Culture Data	√	Not available	Not applicable

Temperature Data: A mean annual ambient temperature was assigned to each subject based on their residential zip code. Temperatures used were 30 year averages of annual temperatures (1981-2010) obtained from The Climate Source (Corvallis, OR). Although recent climate change may have increased these average temperatures for the more recent period in which lung function was collected, it was expected that temperature increases would be similar for all subjects. The use of multiple years of temperature data was employed to minimize year-to-year variation in temperature. These climate data were used for both CF study populations as well as the NHANES study population.

Lung Function Data: Forced expiratory volume in 1 second (FEV_1) is a spirometric measurement obtained at routine and sick visits at CF centers. FEV_1 was used as the measure of lung function for this proposal as it is closely correlated with survival

among CF patients.[31,32] Raw FEV_1 (liters) measurements were not used for this study as both children and adults were included in analyses, thus raw measurements were converted into CF-specific percentiles, which account for age, sex, and body size (height).[50] Measurements obtained before 6 years of age and after lung transplantation were excluded as measurements obtained under the age of 6 years are frequently less reproducible and measurements obtained after transplant reflect lung function of the new non-CF transplanted lungs. The use of CF-specific percentiles also allows for subjects with CF of different ages to be compared to one another despite the decline in lung function seen in CF patients. This phenotype is being utilized as it was the phenotype that has been shown to be associated with ambient temperature and demonstrated heritability.[28,44] In addition, the use of average lung function measures may underestimate lung function in CF patients as CF patients do experience intermittent decreases in lung function associated with respiratory infections; thus, an "average lung function" measure may not capture the true lung function of a subject with CF when well, and the best percentile in the most recent year of data was used to correspond to the most recent location for a subject. The available data that we had for replication from the CF Patient Registry did not include longitudinal measures of lung function. Lastly, a non-CF based percentile for lung function was used for the general population NHANES study population as described in Section 1.5.

Infection Data: Two variables related to each specific infectious organism were collected, namely (1) whether an individual had positive cultures for an organism as a qualitative measure of infection, and (2) the age at which that organism was acquired as quantitative measure. Quantitative measures of the burden of infection, such as colony

counts or cumulative burden of disease, were not available. Subjects were considered to be positive for an organism if they had any respiratory cultures positive for that organism. Age of acquisition for an organism was defined as the date of the first positive culture following at least one prior negative culture (Twin-Sibling Study only).[51] Thirteen CF respiratory pathogens including *Staphylococcus aureus*, Methicillin resistant *S. aureus* (MRSA), non-tuberculous mycobacteria (NTM), *Aspergillus fumigatus*, *Stenotrophomonas maltophilia*, *Achromobacter xylosoxidans*, *Haemophilus influenzae*, *Streptococcus pneumoniae*, *Escherichia coli*, *Klebsiella pneumoniae, Pseudomonas aeruginosa*, mucoid *P. aeruginosa*, and *Burkholderia cepacia* were analyzed.

Section 1.4. Approach for Chapter 3

To identify gene-environment interactions that explain the association between temperature and lung function in (a) known genetic modifiers of CF lung disease, and/or (b) known genetic modifiers of lung function previously identified in the general population.

Introduction: With thousands of genetic variants within the human genome and an equally large number of potential environmental factors that may affect lung function, the risk of false discovery is high when seeking evidence of gene-environment interactions. To minimize this risk, Chapter 3 reduced uncertainty by testing only replicated genetic modifiers of lung function and the replicated environmental factor of ambient temperature. Genetic modifiers tested included both those operating in CF and also those operating on lung function in the general population. The five known CF genetic loci tested included two replicated loci (chromosomes 11p13 and 20q13.2) identified through

genome-wide association and linkage studies by the North American CF Modifier Consortium,[52] and polymorphisms in the mannose binding lectin (*MBL2*) gene supported through meta-analysis,[53] transforming growth factor beta-1 (*TGFB1*) supported through multiple studies,[54-58] and interferon-related developmental regulator 1 (*IFRD1*) identified via genome-wide association and replicated in an independent population.[59] The four known CF genetic loci tested were identified and replicated in large (n=7691 to 20,890) population-based genome-wide association studies of lung function (FEV_1 and FEV_1/FVC ratio) in the general population.[60-63] These genes associated with lung function include hedgehog interacting protein (*HHIP*), advanced glycosylation end product-specific receptor (*AGER*), C-terminal domain containing glutathione S-transferase (*GSTCD*), and 5-hydroxytryptamine receptor 4 (*HTR4*). Together these analyses may provide information on why lung function is highly variable in CF.

Genetic Data: Genotyping has been performed for a sample of the Twin-Sibling Study population using the Illumina 610-Quad Array®, 660W-Quad Array®, and the Omni5 Array® under stringent quality control. Two of the known CF genetic loci have been identified using this array, specifically SNPs rs12793173 and rs4811626. For the other three known CF loci, polymorphisms for a subset of the Twin-Sibling Study have been typed. These include codons 52, 54, and 57 in exon 1 (D, B, and C variant alleles, respectively) and promoter polymorphisms at positions -550 (H/L variant) and -221 (X/Y variant) for *MBL2*, -509 promoter site and codon 10 in *TGFB1*, and the rs7817 polymorphism in *IFRD1*. The known genetic modifiers of lung function in general population examined included *HHIP*, *AGER*, *GSTCD*, and *HTR4*. Analysis of these modifiers was imputation-based as they have not been genotyped in the Twin-Sibling

Population. The correlation between observed and imputed alleles with a similar 660K platform is 0.93, suggesting that imputing a marker for 1000 subjects provides similar data to genotyping that marker for 930 subjects.[64,65] Within the CF Twin-Sibling Study there were a total of 1660 geocoded individuals with lung function and data on at least one polymorphism.

Section 1.5. Approach for Chapter 4

To determine if temperature is associated with lung function in the U.S. general population.

Introduction: Although at a least portion of the association between temperature and lung function (that mediated by *P. aeruginosa*) is specific to individuals with CF, it is conceivable that some of the mechanisms contributing to this association are independent of CF. Thus, a multivariate regression was performed to identify associations between temperature and lung function in the general population.

Study Populations: NHANES: The National Health and Nutrition Examination Survey (NHANES) is an ongoing survey conducted by CDC. It was designed to provide national estimates of the health and nutritional status of the United States' civilian, non-institutionalized population aged two months and older. Surveys are periodically conducted to assess a variety of health outcomes and obtain clinical data, and some of these surveys included lung function testing. The three NHANES surveys that collected lung function data were the NHANES III (1988-1994) survey, the 2007-2008 survey, and the 2009-2010 survey. Lung function from the 2011-2012 survey was not released until December 2014 after analysis had been completed for Chapter 4. The NHANES III

database contains 20,757 individuals with pulmonary function tests obtained between 1988 and 1994 from 89 survey locations across the U. S. Of note the NHANES III population was used to generate one of the most widely used population norms used for comparison of patients in clinical practice in the U.S.[46] The 2007-2008 and 2009-2010 surveys contain 6319 and 7079 subjects with lung function data, respectively.

Lung Function: The clinical standard FEV_1 percentile based on the NHANES population was used,[46] and the best lung function value recorded for a subject was used per clinical practice.

Temperature Data: Mean annual ambient temperature was generated for each individual in a similar manner to Chapters 2 and 3 using residential zip code.

Questionnaire Data: Questionnaire data were used to exclude individuals with active tobacco use or asthma to provide a representative "healthy" study population. Chronic obstructive pulmonary disease (COPD) was not specifically captured by these databases, and thus patients with COPD cannot be excluded.

CHAPTER 2. RESPIRATORY PATHOGENS MEDIATE THE ASSOCIATION BETWEEN LUNG FUNCTION AND TEMPERATURE IN CYSTIC FIBROSIS

Section 2.1. Abstract

Introduction: Mean annual ambient temperature is a replicated environmental modifier of cystic fibrosis (CF) lung disease with warmer temperatures being associated with lower lung function. The mechanism of this relationship is not completely understood. However, *Pseudomonas aeruginosa,* a pathogen that infects the lungs of CF individuals and decreases lung function, also has a higher prevalence in individuals living in warmer climates. We therefore investigated the extent to which respiratory pathogens mediated the association between temperature and lung function.

Methods: Thirteen respiratory pathogens common in CF were assessed in multistep fashion using clustered linear and logistic regression to determine if any mediated the association between temperature and lung function. Analysis was performed in the CF Twin-Sibling Study (n=1730; primary population); key findings were then evaluated in the U.S. CF Foundation Data Registry (n=15,174; replication population).

Results: In the primary population, three respiratory pathogens (*Pseudomonas aeruginosa,* mucoid *P. aeruginosa*, and methicillin-resistant *Staphylococcus aureus*) mediated the association between temperature and lung function. *P. aeruginosa* accounted for 19% of the association (p=0.003), mucoid *P. aeruginosa* for 31% (p=0.001), and MRSA for 13% (p=0.023). The same three pathogens mediated association in the replication population (7%, p<0.001; 7%, p=0.002; 4%, p=0.002, respectively).

Conclusions: We found that three important respiratory pathogens in CF mediated the association between lower lung function and warmer temperatures. These findings have implications for understanding regional variations in clinical outcomes, and interpreting results of epidemiologic studies and clinical trials that encompass regions with different ambient temperatures.

Section 2.2. Introduction

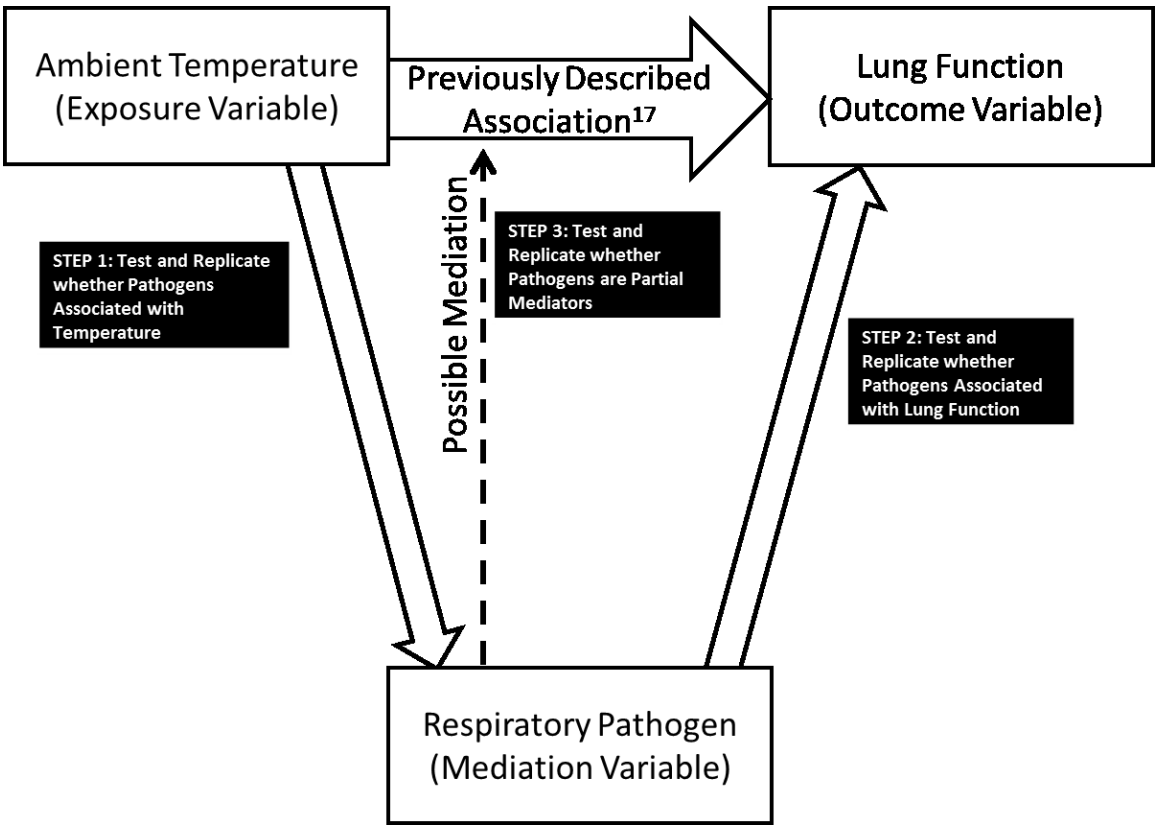
Cystic fibrosis (CF) is a life-limiting autosomal recessive disorder caused by mutations in the *CFTR* gene. Progressive obstructive lung disease and recurrent respiratory infections account for the majority of morbidity and mortality associated with CF. Although over 2000 variants within the *CFTR* gene have been described,[66] individuals with identical mutations who are the same age can have dramatically different severity of lung disease.[3,4,67] Family-based studies suggest that half of this variation is attributable to genetic modifiers other than *CFTR* and half of this variation is due to environmental and/or stochastic (random) factors.[29,30] Even though *CFTR*-specific mutation therapies are now clinically available, these new drugs are extremely expensive.[68,69] Thus, it still remains important to identify specific environmental factors that ameliorate or exacerbate lung disease.

The majority of previous studies examining the role of environmental factors that impact CF lung disease have focused on micro-environmental factors, or those specific to a particular patient or household. Examples of micro-environmental factors demonstrated to impact CF lung disease include secondhand smoke exposure,[33-37] household income,[39,70,71] maternal education,[39] and health insurance status.[39-42,44] Research on the

effects of macro-environmental factors, that is factors that simultaneously affect groups of patients by geographic area, has been more limited with the exception of air pollution.[38]

We previously described an association between climate, specifically mean annual ambient temperature, on lung function in CF where warmer temperature were associated with lower lung function among 3 independent populations of individuals with CF in the continental United States and Australia.[44] Our analyses also suggested that a portion of this association (6-22%) was mediated by a higher prevalence of the common CF respiratory pathogen *Pseudomonas aeruginosa* in warmer climates. *Pseudomonas aeruginosa* was also acquired earlier in warmer regions in retrospectively collected data from the United States and in prospectively collected data from Australia. Subsequently, geographic trends with the first acquisition/prevalence of *P. aeruginosa* were observed in other CF populations,[72] including trends towards earlier acquisition and higher prevalence in the Southern states of the U.S.[73,74] Given that *P. aeruginosa* could have accounted for a fraction of the association between lung function and ambient temperature, we sought to determine if any other common CF respiratory pathogens contributed to the remainder of this variation (See Figure 2.1 for graphical depiction of study model).

Figure 2.1. Graphical Depiction of Mediation Scheme



Section 2.3. Methods

Ethics Statement and Recruitment: Written informed consent was obtained from participants (or their parents/legal guardians) enrolled in the CF Twin-Sibling Study (PI: G.R. Cutting). This study, including the CF Foundation Data Registry data downloads, was approved by the Johns Hopkins University Institutional Review Board. Primary Population: U.S. CF Twin and Sibling Study (CFTSS): Participants were recruited from CF centers based on having a twin or sibling also with CF (n=2086 in 1018 families).[28] Longitudinal data with multiple assessment timepoints were collected between 10/27/00

and 3/31/13 with data supplementation from the U.S. CF Foundation Data Registry through 12/31/11. Replication Population: U.S. CF Foundation Patient Registry (CFFDR): Anonymized cross-sectional data from the calendar year 2007 was provided (n=24,799). Participants were excluded if lung function, respiratory culture, or residential postal/zip code data were unavailable or they were living outside of the continental United States, or if self-reported to be actively smoking. CFTSS participants that were known to be enrolled in the CFF Patient Registry (n=1435) were excluded from the CFF sample.

Exposure (Temperature) Variable: U.S. participants were mapped to the center of their most recent known residential zip code (finest resolution available) to derive measures for mean annual ambient temperature as this measure of temperature was previously found to be associated with lung function. Source data for mapping was 1981-2010 mean annual ambient temperatures (The Climate Source; Corvallis, OR; 400 meter resolution).

Outcome (Lung Function) Variable: Raw FEV_1 (liters) measurements were converted into CF-specific percentiles excluding measurements obtained before 6 years of age and after lung transplantation.[50] Lung function was defined as the best percentile in the most recent year of data.

Potential Mediation (Respiratory Pathogen) Variable: Using all available respiratory culture data, participants were considered to be positive for a common CF respiratory pathogen (*Achromobacter xylosoxidans, Aspergillus spp., Burkholderia cepacia* complex, *Escherichia coli, Haemophilus influenzae, Klebsiella pneumoniae,*

non-tuberculous *Mycobacteria spp., Pseudomonas aeruginosa,* mucoid *P. aeruginosa, Staphylococcus aureus,* Methicillin-resistant *Staphylococcus aureus, Stenotrophomonas maltophilia,* and/or *Streptococcus pneumoniae*) if they had any cultures positive for that pathogen. Data on culture source were not available for this study, but typically most clinical CF cultures were obtained via throat swab or expectorated sputum.

Data Analysis: CFTSS participants served as the primary population and CFFDR as the replication population as the CFTSS participants were better characterized and had longitudinal data. Regressions clustered by family (CFTSS) or by CF care center (CFFDR), chi square, and student's t-tests were performed using Stata 11 (StataCorp LP; College Station, TX). To establish whether the presence of individual respiratory pathogens (mediation variable) mediate the association between ambient temperature (exposure variable) and lung function (outcome variable), a standardized multistep process was utilized (See Figure 2.1 for schematic).[48] The initial step of establishing the relationship between exposure and outcome was previously demonstrated in both the primary and replication populations.[44]

The first step for this study was to determine whether mediation variables were associated with exposure through Generalized Estimating Equations (clustered regression) with a specific pathogen as the dependent variable, temperature as the independent variable, and adjustment for age. Using only mediation variables that were associated with exposure, the second step was to determine whether the mediation variables were associated with outcome through clustered linear regression with lung function as the dependent variable, a specific pathogen as an independent variable, and adjusted for age. Using only mediation variables that were independently associated with

exposure and outcome, the third step was to determine whether the mediation variables altered the relationship between the exposure and outcome variables through clustered linear regression with lung function as the dependent variable, a specific pathogen and temperature as independent variables; these regressions were also adjusted for age at the time of lung function and health insurance status based on previously published models.[44] We conducted these analyses separately in the primary and replication populations.

The effect of mediation by infection on lung function was estimated by dividing the difference of regression coefficients of temperature adjusted and unadjusted for the mediator by coefficient of temperature unadjusted for the mediator.[48,49] Confidence intervals for mediation were derived using the mediation package in STATA with clustered standard errors.[75] *P* values for mediation were derived via the Sobel test.[76,77]

Section 2.4. Results

Demographics: The primary population for this study consisted of 1730 participants from the CF Twin-Sibling Study (CFTSS)(n=2086) and the replication population for this study consisted of 15,174 participants from the U.S. CF Foundation Data Registry (CFFPR)(n=24,799) after exclusions. Comparisons of included vs. excluded participants are detailed in Tables 2.1 and 2.2.

Table 2.1. Primary Study Population (CF Twin-Sibling Study) Included and Excluded Participants

Mean ± S.D. [Range]		**Entire CF Twin-Sibling Study** (n = 2086)	**Included Participants** (n = 1730)	**Excluded Participants** (n = 356)	***P* value**
Demographics	**Sex** (% male)	52.7%	51.7%	57.9%	**0.033**
	Race/Ethnicity (% nonwhite)	9.4% (n = 2083)	9.7% (n = 1729)	8.2% (n = 354)	0.39
	Age at Diagnosis (years)	2.2 ± 5.3 [0, 52] (n = 2084)	2.3 ± 5.4 [0, 52]	1.6 ± 4.3 [0, 33] (n = 354)	**0.010**
	***CFTR F508del* Mutations** (% homozygous)	48.1% (n = 2078)	48.6% (n = 1728)	45.4% (n = 350)	0.28
Clinical Data	**Age at Last Respiratory Culture** (years)	18.1 ± 10.3 [0, 67.0] (n = 2056)	19.6 ± 9.7 [5.3, 67.0]	10.4 ± 9.6 [0, 47.4] (n = 326)	**<0.001**
	Age at FEV_1 Measurement (years)	18.9 ± 9.6 [6, 63.9] (n = 1894)	19.0 ± 9.6 [6, 63.9]	18.2 ± 9.5 [6, 48.4] (n = 164)	0.32
	CF-Specific FEV_1 (percentile)	69.2 ± 27.2 [0, 100] (n = 1894)	69.5 ± 26.9 [0, 100]	66.9 ± 30.5 [0, 100] (n = 164)	0.26
Exposures	**Mean Annual Ambient Temperature** (°F)	54.3 ± 7.1 [38.5, 76.5] (n = 1901)	54.3 ± 7.1 [38.5, 76.5]	54.3 ± 7.1 [42.6, 72.5] (n = 171)	0.91
	Public Insurance (% yes)	32.2% (n = 1933)	31.9%	35.0% (n = 203)	0.37

Table 2.2. Replication Study Population (CF Foundation Data Registry) Included and Excluded Participants

Mean ± S.D. [Range]		**Entire CF Foundation Data Registry** (n = 24,799)	**Included Participants** (n = 15,174)	**Excluded Participants** (n = 9625)	***P* value**
Demographics	**Sex** (% male)	51.9%	52.1%	51.7%	0.57
	Race/Ethnicity (% nonwhite)	12.1% (n = 24,721)	11.2% (n = 15,147)	13.6% (n = 9574)	**<0.001**
	Age at Diagnosis (years)	3.5 ± 7.9 [0, 78.5]	3.8 ± 8.2 [0, 73.7]	2.8 ± 7.4 [0, 78.5]	**<0.001**
	***CFTR F508del* Mutations** (% homozygous)	49.0% (n = 22,248)	49.8% (n = 13,644)	47.7% (n = 8604)	**0.002**
Clinical Data	**Age at Last Respiratory Culture** (years)	17.7 ± 12.2 [0, 74.2] (n = 22,801)	20.7 ± 11.4 [5.7, 74.2]	11.7 ± 11.6 [0.1, 73.3] (n = 7627)	**<0.001**
	Age at FEV_1 Measurement (years)	20.8 ± 11.5 [6.0, 79.3] (n = 19,252)	20.5 ± 11.4 [6.0, 74.0]	22.1 ± 11.8 [6.0, 79.3] (n = 4078)	**<0.001**
	CF-Specific FEV_1 (percentile)	65.9 ± 26.3 [0, 100] (n = 19,252)	65.3 ± 26.4 [0, 100]	67.8 ± 25.9 [0, 100] (n = 4078)	**<0.001**
Exposures	**Mean Annual Ambient Temperature** (°F)	55.3 ± 7.7 [32.2, 77.5] (n = 22,940)	55.3 ± 7.8 [34.3, 77.5]	55.2 ± 7.6 [32.2, 76.3] (n = 7766)	0.36
	Public Insurance (% yes)	42.0% (n = 23,697)	40.5%	44.5% (n = 8523)	**<0.001**

In both populations, excluded participants had an earlier age of diagnosis and younger age at last respiratory culture. In the primary CFTSS population, excluded participants were also more likely to be male, and in the replication CFFDR population, excluded participants were more likely to be non-white, older at the measurement of lung function, have a higher CF-specific FEV_1, and less likely to be homozygous for the most common *CFTR* mutation (*F508del*).

There were also significant differences between the primary CFTSS and the replication CFFDR populations (Table 2.3), but owing to the large sample size of the

CFF population (n=15,174) and the magnitude of the differences, most differences were not clinically relevant.[78] Still, a few differences were notable.

Table 2.3. Primary and Replication Study Population Demographics

	Mean ± S.D. [Range]	**CF Twin-Sibling Study (Primary)** (n = 1730)	**CF Foundation Data Registry (Replication)** (n = 15,174)	***P* Value**
Demographics	**Sex** (% male)	51.7%	52.1%	0.75
	Race/Ethnicity (% nonwhite)	9.7% (n = 1729)	11.2% (n = 15,147)	**0.048**
	Age at Diagnosis (years)	2.3 ± 5.4 [0, 52]	3.8 ± 8.2 [0, 73.7]	**<0.001**
	***CFTR F508del* Mutations** (% homozygous)	48.6% (n = 1728)	49.8% (n = 13,644)	0.34
Clinical Data	**Age at Last Respiratory Culture** (years)	19.6 ± 9.7 [5.3, 67.0]	20.7 ± 11.4 [5.7, 74.2]	**<0.001**
	Age at FEV_1 Measurement (years)	19.0 ± 9.6 [6, 63.9]	20.5 ± 11.4 [6.0, 74.0]	**<0.001**
	CF-Specific FEV_1 (percentile)	69.5 ± 26.9 [0, 100]	65.3 ± 26.4 [0, 100]	**<0.001**
Exposures	**Mean Annual Ambient Temperature** (°F)	54.3 ± 7.1 [38.5, 76.5]	55.3 ± 7.8 [34.3, 77.5]	**<0.001**
	Public Insurance (% yes)	31.9%	40.5%	**<0.001**

The primary CFTSS population had a younger age of diagnosis (2.3 vs. 3.8 years) and a higher CF-specific FEV_1 measurements (69.5 vs. 65.3) than the replication CFFDR population. Individuals in the primary CFTSS population were less likely to have public health insurance (31.9 vs. 40.5%) than the replication CFFDR population. Some of these differences reflect the recruitment criteria for the CFTSS, namely having 2 living siblings or twins with CF, which resulted in study population that was younger and healthier than the general CF population,[28] as well as earlier diagnosis for at least the younger sibling

owing to familial knowledge. There were no significant differences in the gender composition or frequency of *F508del* homozygotes between the two populations.

The percentage of participants who had a culture with a specific respiratory pathogen ever is described in Table 2.4. For the primary CFTSS population, the percentages ranged from a low frequency of 8.6% for non-tuberculous mycobacteria to a high frequency of 98.0% for *S. aureus*. For the replication CFFDR population, the percentages ranged from 1.3% for *E. coli* and *K. pneumoniae* to 60.4% for *P. aeruginosa.* The prevalence frequencies for any given organism were higher in the CFTSS population compared to the CFFDR population. This pattern of findings likely reflects differences in study design; the CFTSS population incorporated longitudinal data with each subject having 16.2±7.1 years [Range: 0.4, 48.6] of respiratory culture data, whereas the CFFDR population had cross-sectional data with a single calendar year of data.

Table 2.4. Prevalence of Respiratory Pathogens in the Primary and Replication Study Populations

Mean ± S.D. [Range in years]	**Positive on any Respiratory Culture: CF Twin-Sibling** (%; n = 1730 unless otherwise noted)	**Age at First Positive Culture: CF Twin-Sibling** (years)	**Positive on any Respiratory Culture: CF Foundation** (%; n = 15,174)
Achromobacter xylosoxidans	22.9	13.7 ± 9.2 [0.1, 62.9] (n = 389)	6.7
Aspergillus spp.	48.8	14.1 ± 8.5 [0.1, 56.9] (n = 835)	16.4
***Burkholderia cepacia* Complex**	8.8	14.3 ± 8.3 [0.7, 50.4] (n = 152)	3.3
Escherichia coli	13.7 (n = 1258)	3.8 ± 7.5 [0.0, 57.2] (n = 147)	1.3
Haemophilus influenzae	76.5	6.8 ± 7.3 [0.0, 55.0] (n = 1177)	13.7
Klebsiella pneumoniae	9.7	5.9 ± 10.4 [0.0, 61.3] (n = 147)	1.3
Non-tuberculous *Mycobacteria spp.*	8.6	19.6 ± 8.9 [4.9, 49.7] (n = 148)	2.5
Pseudomonas aeruginosa	89.5	6.8 ± 6.8 [0.0, 53.8] (n = 1286)	60.4
Mucoid *Pseudomonas aeruginosa*	62.7	11.7 ± 8.4 [0.2, 53.8] (n = 1030)	42.6
Staphylococcus aureus	98.0	6.2 ± 7.2 [0.0, 54.6] (n = 1227)	50.5
Methicillin-resistant *Staphylococcus aureus*	51.6	13.3 ± 8.9 [0.0, 58.9] (n = 863)	22.9
Stenotrophomonas maltophilia	50.8	11.8 ± 9.0 [0.0, 60.2] (n = 854)	13.4
Streptococcus pneumoniae	42.8 (n = 1258)	7.8 ± 7.6 [0.0, 44.9] (n = 511)	Data not available

Mediation STEP 1: Determining whether the mediation variable is associated with exposure: As previously mentioned, having previously established the association between temperature and lung function, we used a three step process to establish whether a specific infectious pathogen mediated the association between temperature and lung function (Figure 2.2).

To assess whether annual ambient temperature was associated with specific respiratory pathogens, logistic regression was conducted in the primary CFTSS population for each pathogen. Clustering by family (adjusted estimates of variance) was performed as the presence of an organism for a subject may be influenced by his/her sibling/twin's potential colonization with specific respiratory pathogens (Table 2.5). Of the 13 pathogens assessed, seven were associated with annual ambient temperature in the primary population, namely *A. xylosidans*, *H. influenzae*, *K. pneumoniae*, non-tuberculous mycobacteria*, P. aeruginosa*, mucoid *P. aeruginosa*, and MRSA. Of these 7 pathogens, only five were associated with temperature in the CFFDR replication population, namely *H. influenzae*, non-tuberculous mycobacteria*, P. aeruginosa*, mucoid *P. aeruginosa*, and MRSA. Of note, there was a higher odds ratio for the presence of a pathogen associated with a warmer mean annual ambient temperature for these pathogens, excepting *H. influenzae* whose presence was associated with colder temperatures.

Table 2.5. Mediation STEP 1: Testing for the Association of Exposure Variable (Temperature) and Potential Mediation Variable (Respiratory Pathogen) in Primary and Replication Populations

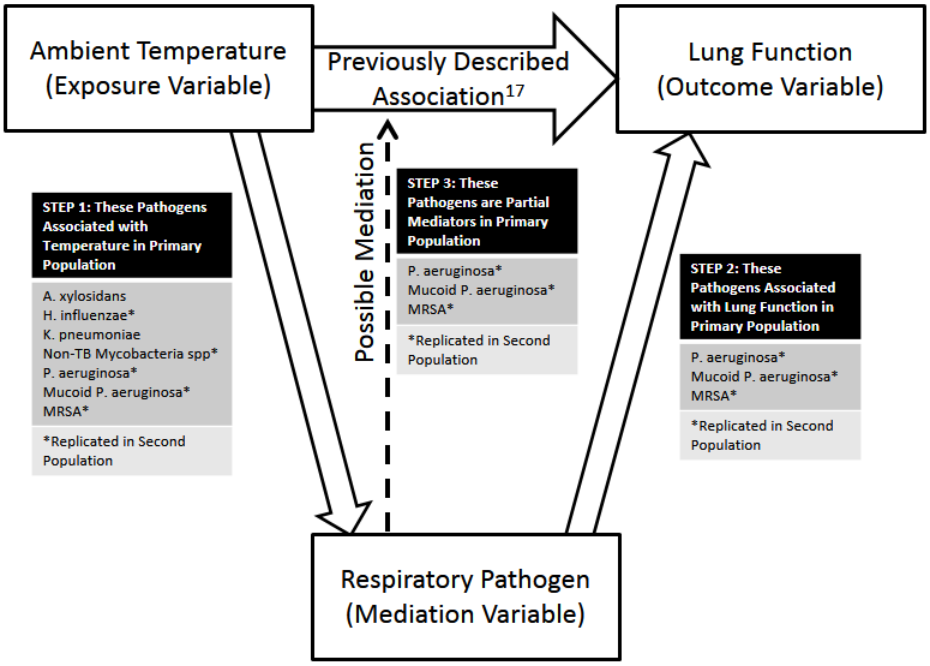
Co-efficient [95%CI] (*p* value)[1]	**CF Twin-Sibling Study (Primary; n = 1730 unless otherwise specified)**		**CF Foundation Data Registry (Replication; n = 15,174)**	
	Model *p* Value	**Temperature Odds Ratio[2]** (per 10°F)	**Model *p* Value**	**Temperature Odds Ratio[3]** (per 10°F)
Achromobacter xylosoxidans	<0.001	1.27 [1.06, 1.53] **(0.010)**	0.011	0.99 [0.86, 1.13] (0.85)
Aspergillus spp.	<0.001	1.16 [0.99, 1.35] (0.06)	-	-
***Burkholderia cepacia* Complex**	<0.001	1.02 [0.78, 1.34] (0.87)	-	-
Escherichia coli	0.04 (n =1258)	0.77 [0.58, 1.03] (0.08)	-	-
Haemophilus influenzae	<0.001	0.72 [0.60, 0.86] **(<0.001)**	<0.001	0.77 [0.62, 0.97] **(0.023)**
Klebsiella pneumoniae	0.04	0.73 [0.55, 0.97] **(0.031)**	0.36	0.89 [0.67, 1.18] (0.41)
Non-tuberculous *Mycobacteria spp.*	<0.001	1.82 [1.42, 2.33] **(<0.001)**	<0.001	1.80 [1.50, 2.17] **(<0.001)**
Pseudomonas aeruginosa	<0.001	1.53 [1.14, 2.05] **(0.004)**	<0.001	1.17 [1.09, 1.25] **(<0.001)**
Mucoid *Pseudomonas aeruginosa*	<0.001	1.34 [1.13, 1.60] **(0.001)**	<0.001	1.14 [1.06, 1.24] **(0.001)**
Staphylococcus aureus	0.13	0.85 [0.55, 1.33] (0.48)	-	-
Methicillin-resistant *Staphylococcus aureus*	<0.001	1.48 [1.25, 1.75] **(<0.001)**	<0.001	1.17 [1.07, 1.27] **(0.001)**
Stenotrophomonas maltophilia	0.82	0.97 [0.83, 1.12] (0.64)	-	-
Streptococcus pneumoniae	0.04 (n =1258)	0.82 [0.67, 1.01] (0.06)	-	-

[1]All logistic regressions were clustered by family for the primary population and CF center for the replication population and adjusted for age at the time of the last respiratory culture in years. Regressions for the replication population were only performed for a respiratory pathogen if the pathogen was associated with temperature in the primary population.

[2]Temperatures odds ratios represents the adjusted odds ratio that an individual has a positive culture for the specified respiratory pathogen for a 10°F increase in mean annual ambient temperature. I.e., there is a 1.27 increased odds of having a positive culture for *A. xylosoxidans* in the primary population for every 10°F increase in the mean annual ambient temperature.

[3]Temperature odds ratios were not calculated for the replication population if the odds ratio for the primary population was non-significant ($p>0.05$).

Figure 2.2. Graphical Depiction of Mediation Results



Mediation STEP 2: Determining whether the mediation variable is associated with outcome: To assess whether specific respiratory pathogens were associated with lung function, clustered linear regression was conducted in the primary CFTSS population for the pathogens identified in Step 1 (Table 2.6). Of the 5 pathogens assessed, only 3 were associated with lung function (CF-specific FEV_1) in the primary population, namely *P. aeruginosa*, mucoid *P. aeruginosa*, and MRSA. All three of these pathogens were also associated with lung function in the CFFDR replication population. Each of these three pathogens was associated with a reduction in lung function in both populations.

Table 2.6. Mediation STEP 2: Testing for the Association of Potential Mediation Variable (Respiratory Pathogen) and Outcome Variable (Lung Function) in Primary and Replication Populations

Co-efficient [95%CI] (*p* value)[1]	**CF Twin-Sibling Study (Primary; n = 1730)**		**CF Foundation Data Registry (Replication; n = 15,174)**	
	Model *p* Value	**Change in Lung Function[2] (with specified respiratory pathogen)**	**Model *p* Value**	**Change in Lung Function[3] (with specified respiratory pathogen)**
Haemophilus influenzae	0.004	1.9 [-1.4, 5.2] (0.26)	-	-
Non-tuberculous *Mycobacteria spp.*	0.006	-1.8 [-6.6, 3.1] (0.48)	-	-
Pseudomonas aeruginosa	<0.001	-14.8 [-18.0, -11.5] **(<0.001)**	<0.001	-7.5 [-8.6, -6.5] **(<0.001)**
Mucoid *Pseudomonas aeruginosa*	<0.001	-13.3 [-16.1, -10.6] **(<0.001)**	<0.001	-8.3 [-9.4, -7.2] **(<0.001)**
Methicillin-resistant *Staphylococcus aureus*	<0.001	-4.6 [-7.4, -1.9] **(0.001)**	<0.001	-5.3 [-6.5, -4.2] **(<0.001)**

[1]All logistic regressions were clustered by family for the primary population and CF center for the replication population and adjusted for age at the time of pulmonary function testing. Regressions for the primary population were only performed for a respiratory pathogen if the pathogen was associated with temperature in the replication population (Table 2.5). Regressions for the replication population were only performed for a respiratory pathogen if the pathogen was associated with lung function in the primary population.

[2]Change in lung function coefficients represent the change in CF-specific FEV_1 percentile that an individual with a positive culture for the specified respiratory pathogen has compared to an individual with negative cultures for that pathogen. I.e., the CF-specific FEV1 percentile is 14.8 percentile points lower on average in the primary population with at least one positive respiratory culture for *P. aeruginosa*.

[3]Change in lung function co-efficients were not calculated for the replication population if the change for the primary population was non-significant ($p>0.05$).

Mediation STEP 3: Determining whether the mediation variable is associated with outcome: To assess whether specific infectious organisms mediate the association between ambient temperature and lung function, we compared the effect of temperature on lung function in a base model without a respiratory pathogen to a mediation model which included one of the 3 respiratory pathogens identified in Step 2 (Table 2.7). All 3 of the organisms identified in Step 2 partially mediated the association between temperature and lung function in both the primary and the replication populations. In the

primary CFTSS population, *P. aeruginosa* was found to account for 19% of the association (p=0.003; 95% C.I.: 10%, 79%) between temperature and lung function, mucoid *P. aeruginosa* was found to account for 31% (p=0.001; 95% C.I.: 16%, 100%), and MRSA for 13% (p=0.023; 95% C.I.: 7%, 60%). The degree of mediation observed in the replication CFFDR population was lower overall, with *P. aeruginosa* accounting for 7% of the association (p<0.001; 95% C.I.: 6%, 11%) between temperature and lung function, mucoid *P. aeruginosa* accounting for 7% (p=0.002; 95% C.I.: 5%, 10%), and MRSA for 4% (p=0.002; 95% C.I.: 3%, 5%). We also assessed the effect of all three of these organisms simultaneously in both the CFTSS and CFFDR populations by placing all 3 organisms into the base model (adjusted for insurance status and age). The inclusion of all 3 organisms in the model suggested that these infectious organisms account for 43% of the association between temperature and lung function in the primary CFTSS population, and 12% in the replication CFFDR population.

Table 2.7. Mediation STEP 3: Testing whether the Mediation Variable (Respiratory Pathogen) Mediates the Association between the Exposure Variable (Temperature) and the Outcome Variable (Lung Function) in Primary and Replication Populations

Co-efficient [95%CI] (*p* value)[1]	**CF Twin-Sibling Study (Primary; n = 1730)**			**CF Foundation Data Registry (Replication; n = 15,174)**		
	Change in Lung Function (per 10°F)	**Change in Lung Function** (with specified respiratory pathogen)	**Percent Mediation[2]**	**Change in Lung Function** (per 10°F)	**Change in Lung Function** (with specified respiratory pathogen)	**Percent Mediation**
Base Model (without respiratory pathogen)	-2.6 [-4.8, -0.4] **(0.021)**	-	-	-3.1 [-4.1, -2.1] **(<0.001)**	-	-
***Pseudomonas aeruginosa* Model**	-2.1 [-4.3, 0.1] (0.06)	-13.5 [-16.8, -10.3] **(<0.001)**	19% [10%, 79%] **(0.003)**	-2.9 [-3.9, -1.9] **(<0.001)**	-6.9 [-8.0, -5.9] **(<0.001)**	7% [6%, 11%] **(<0.001)**
Mucoid *Pseudomonas aeruginosa* Model	-1.8 [-3.9, 0.3] (0.10)	-12.9 [-15.6, -10.2] **(<0.001)**	31% [16%, 100%] **(0.001)**	-2.9 [-3.9, -1.9] **(<0.001)**	-7.5 [-8.6, -6.5] **(<0.001)**	7% [5%, 10%] **(0.002)**
Methicillin-resistant *Staphylococcus aureus* Model	-2.3 [-4.5, 0.0] (0.05)	-3.6 [-6.4, -0.9] **(0.010)**	13% [7%, 60%] **(0.023)**	-3.0 [-4.0, -2.0] **(<0.001)**	-4.2 [-5.3, -3.1] **(<0.001)**	4% [3%, 5%] **(0.002)**
Combined Model with mucoid and non-mucoid *P. aeruginosa* and MRSA	-1.5 (-3.7, 0.7) (0.18)	P. aeruginosa -6.3 **(<0.001)** Mucoid *P. aeruginosa* -10.8 **(<0.001)** MRSA -2.5 **(0.001)**	43%	-2.7 (-3.7, -1.7) **(<0.001)**	P. aeruginosa -3.2 **(<0.001)** Mucoid *P. aeruginosa* -5.3 **(<0.001)** MRSA -3.9 **(<0.001)**	12%

[1]All logistic regressions were clustered by family for the primary population and CF center for the replication population and adjusted for age at the time of pulmonary function testing and health insurance status. Regressions for the primary population were only performed for a respiratory pathogen if the pathogen was associated with lung function in the replication population (Table 2.6). Change in lung function coefficients represent the change in CF-specific FEV_1 percentile for a 10°F increase in mean annual ambient temperature as well as a change in an individual with a positive culture for the specified respiratory pathogen has compared to an individual with negative cultures for that pathogen.

[2]Percent mediation is derived by dividing the temperature coefficient from a model with a respiratory pathogen by the temperature coefficient from the base model, then subtracting that from 1. Confidence intervals and Sobel test p-values for mediation were obtained as described in the Methods and are not available for models with more than one mediator (i.e., more than one infectious organism).

Section 2.5. Discussion

A number of human pathogens are affected by climate, most notably those with arthropod vectors, but also water- and food-borne pathogens.[45] Our study suggests that

pulmonary bacterial infections or colonization are associated with climatic factors in individuals with a chronic respiratory disease.[44] We previously reported an association between mean annual ambient temperature and lung function in individuals with CF. Among 13 CF respiratory pathogens, we found that *P. aeruginosa* (all strains as well as mucoid only) and MRSA infection mediated the association between lung function and ambient temperature in two independent populations of individuals with CF. This study extends our previous observation[44] that a higher prevalence of *P. aeruginosa* is associated with warmer ambient temperature, and indicates that climactic factors may explain the higher rates of *P. aeruginosa* and MRSA lung infection in CF noted in Southern regions of the United States.[74]

The factors behind the geographic distribution of these respiratory pathogens in CF are not clear. Possibilities include factors that favor acquisition or colonization of individuals with CF in warmer climates, a geographic distribution that mirrors that of the general population, and/or distribution by chance alone. Although *P. aeruginosa* is not a common pathogen in the general population, MRSA is. Limited data on the proportion of MRSA isolates among all *S. aureus* isolates suggests that there may be geographical trends in the distribution of the proportion of MRSA isolates from *S. aureus* isolates in the United States and Europe (Figures 2.3 and 2.4). Alternatively, warmer temperatures may favor colonization of individuals with CF with respiratory pathogens. For example, *in vitro* studies of biofilm resistance of *P. aeruginosa* and *S. aureus* can increase with increases in incubation temperature under the right circumstances.[79,80] In addition, the adhesion properties of *P. aeruginosa* to surfaces may also vary with temperature.[81]

Figure 2.3. Proportion of MRSA from *S.* aureus isolates in the general population by region in the United States. Map and Data are from The Center of Disease Dynamics, Economics & Policy accessed from their website on 01/27/2015. http://www.cddep.org/projects/resistance_map/community_and_hospital_associated_mrsa

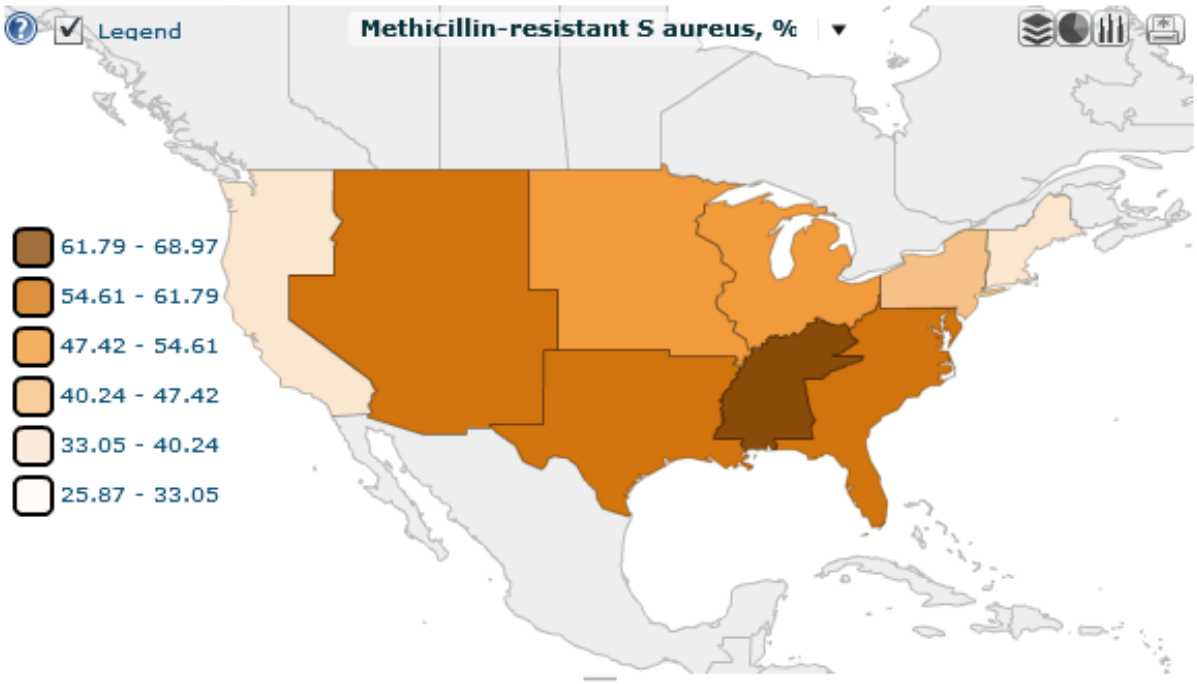
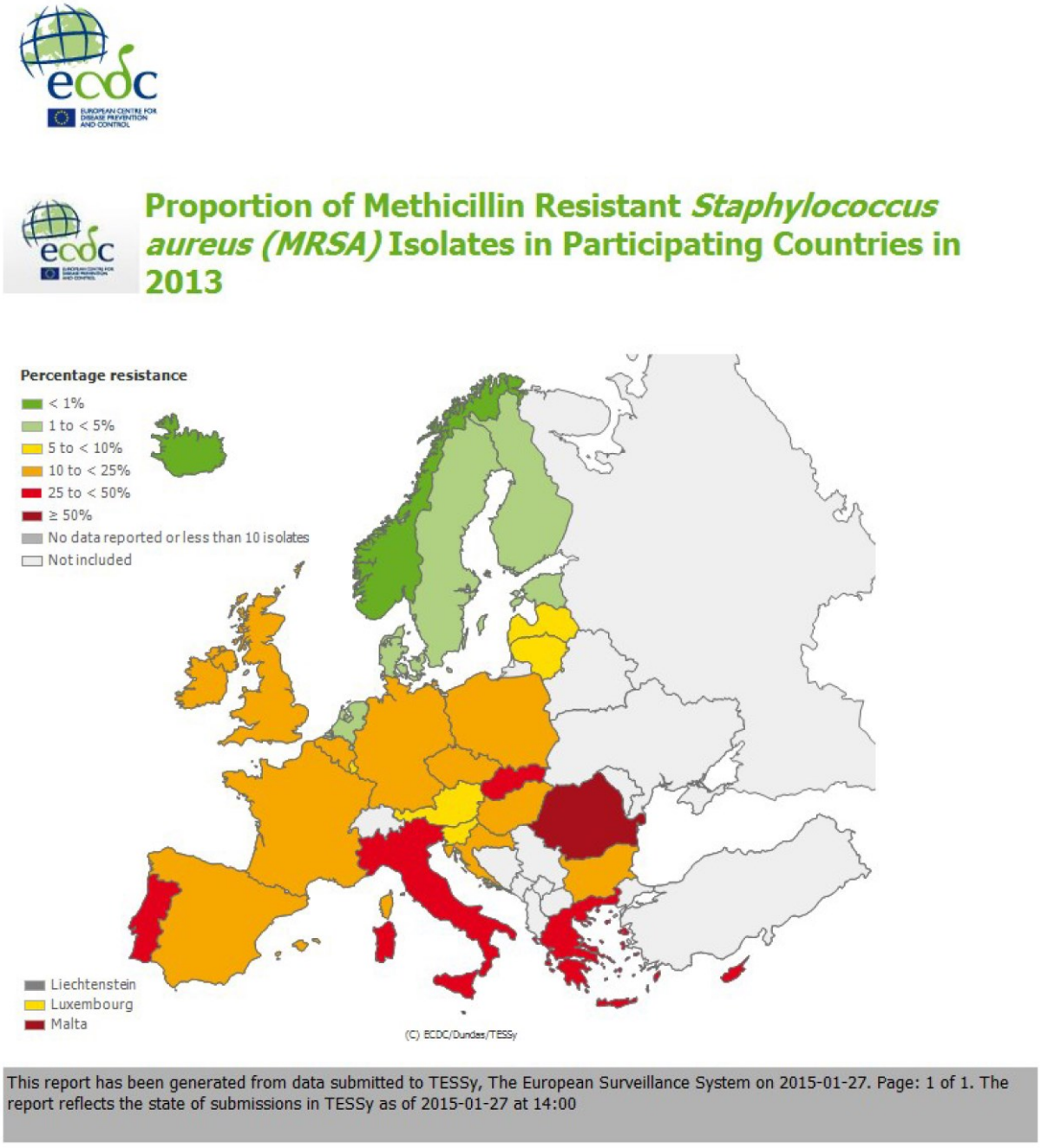


Figure 2.4. Proportion of MRSA from *S.* aureus isolates in the general population by country in Europe. Map and Data are from the European Centers for Disease Prevention and Control accessed from their website on 01/27/2015. http://www.ecdc.europa.eu/en/healthtopics/antimicrobial_resistance/database/Pages/map_reports.aspx



We also found that the portion of the association between temperature and lung function that was mediated by these three pathogens varied by study population, with 43% of the association accounted for by these organisms in the primary CFTSS population, but only 12% in the CFFDR population. The presence of respiratory pathogens was assessed through longitudinal data in the CFTSS population as opposed to cross-sectional data in the CFFDR population; these longitudinal data may provide more accurate information regarding infection status. Alternatively, the CFTSS had higher lung function and was younger than the CFFDR population, so it is possible that the presence of these pathogens had a more significant impact in absolute changes in lung function in this population.

Our study has several strengths, including a well characterized primary population with longitudinal data. Our replication population includes the largest available database in term of numbers of participants with CF. Lung function testing in patients with CF is generally robust with multiple tests performed annually with routine clinical care. Our study also has several limitations, most notably the possibility of unmeasured confounders. For example, *P. aeruginosa* has been characterized as a highly diverse species,[82] and geographic variation in MRSA strains has been documented.[83] Thus it is possible that more virulent strains are geographically distributed in warmer climates by chance alone, and we did not account for strain differences. Likewise, the presence of non-tuberculous mycobacteria in CF respiratory cultures has been associated with unmeasured factor of vapor pressure, but not ambient temperature.[84] Additionally, our model assumed that the mediation variables of infection act upon the outcome variable of lung function, whereas it is also possible that more severe lung disease could lead to an

increased propensity for infections. It is also possible that individuals with CF may have polymorphisms in modifier genes affecting infection acquisition/colonization that are geographically distributed owing to founder effects or immigration patterns. Our study was also subject to misclassification bias by geographic location as participants move over the course of their lifetimes. Still, 79.5% of the CFTSS participants and 77.6% of the CFFDR participants live within the state in which they were born with another 8% (CFTSS) and 7.3% (CFF) living within an adjacent state. Also, our respiratory culture data were obtained from clinical specimens largely obtained through sputum samples and throat swabs; the latter, in particular, may not accurately represent lower airway flora.[18] However, poor quality of culture data would likely bias our results to the null, rather than leading to positive associations. Additionally, although we ascribed observed results to partial mediation it is possible that unmeasured confounders exist or that collider stratification bias was present. Lastly, our use of retrospective culture data may under- or overestimate the prevalence of respiratory pathogens depending on reporting biases by CF care centers who may not report all culture data to the CF Foundation.

Our results imply that where patients live may impact outcomes, in so much as patients living in warmer climates may acquire *P. aeruginosa*, mucoid *P. aeruginosa*, or MRSA at earlier ages and/or it may be more difficult to eradicate or treat these respiratory pathogens in warmer climates. This would suggest that surveillance via respiratory cultures may need to be more frequent, or eradication and prevention may need to be more aggressive in clinics located in warmer climates. The U.S. CF Foundation also compares CF care centers' clinical outcomes (e.g., lung function). Accordingly, comparisons of centers may need to be weighted based on geographic

location as there may be factors dictating respiratory pathogen prevalence and lung function that are macro-environmental factors outside of their control.

In addition, our results also may have implications for clinical studies that are conducted across a diverse geographic area, as results may need to be adjusted for climatic variation. Also, we note that climate change may affect the geographic distribution of specific infectious diseases.[85] Given our observations that warmer temperatures were associated with several respiratory pathogens, it is possible that global warming may lead to a higher prevalence of specific pathogens among individuals with CF. Lastly, while CF patients may view warmer climates as more healthy, it is quite possible that living in warmer climates might lead to worse outcomes than living in colder climates.

In conclusion, the prevalence of three important common respiratory pathogens in CF appeared to mediate the association between lower lung function and warmer temperatures. These findings have implications for understanding regional variations in clinical outcomes, and results of epidemiologic studies and clinical trials which encompass regions with different ambient temperatures.

CHAPTER 3. GENE MODIFIER TEMPERATURE INTERACTIONS IN CYSTIC FIBROSIS

Section 3.1. Abstract

Introduction: Patients with cystic fibrosis (CF) with identical mutations in the disease-causing gene can have drastically different lung function. Twin studies in CF have suggested that approximately 50% of the variation seen in CF lung disease is related to modifier genes and 50% is due to environmental factors. One replicated environmental modifier of CF lung function is mean annual ambient temperature whereby warmer temperatures are associated with lower lung function in CF, but the mechanism is unknown. We sought to test for genetic factors that modify the association of temperature with lung function (gene-environment interactions) that might account for this association.

Methods: Using the residential postal code of participants in the CF Twin-Sibling Study (n=2086), these individuals were geocoded with a mean annual ambient temperature. Multivariate regression was used to test for gene-temperature interactions with 10 known gene modifiers of lung function in CF and 8 known gene modifiers of lung function among the general population, separately.

Results: After Bonferroni correction for 11 independent tests, no significant gene-environment interactions were identified to be associated with lung function (FEV_1) in the CF Twin-Sibling population (all p values >0.0045 [the statistical threshold for significance).

Conclusions: Although we did not identify any gene-environment interactions in our study, it is still remains important to identify gene-environment interactions in CF as well as elucidate the relationship between temperature and lung function in CF to identify susceptible individuals who could benefit from specific interventions to improve longevity.

Section 3.2. Introduction

Cystic fibrosis (CF) is one of the most common autosomal recessive disorders in the United States. The majority of morbidity and mortality associated with CF is due recurrent respiratory infections leading to decline in pulmonary function. Although the gene responsible for CF, *CFTR*, was identified in 1989,[1] and over 1900 genetic variants have been described to date,[66] many of the phenotypic manifestations of CF do not correlate with specific mutations.[86] Most significantly, patients with identical *CFTR* mutations often have drastically different lung function.[2-4] More recent twin and sibling studies in CF have suggested that approximately 50% of the variation seen in CF lung disease is related to modifier genes, and 50% is due to common/unique environmental factors.[29,30] Understanding the sources of variation in CF lung disease is crucial in ameliorating its burden.

Although several genetic modifiers have been identified through genome-wide studies and candidate approaches, the identified loci may not explain the variation seen in CF lung function. A recent GWAS study's identified loci only accounted for 4-46% of the variation in lung function.[52] Likewise, although a number of environmental factors have been associated with variation in CF lung function, based on a multi-factor study,[44]

it is likely that the currently identified environmental factors also only account for a small portion of variation. One previously identified environmental modifier of CF lung function is mean annual ambient temperature whereby warmer temperatures are associated with lower lung function in CF, which we previously described in 3 independent populations of patients with CF.[44]

Unaccounted for sources of variation may be unstudied genetic or environmental factors, but may also include gene-environment interactions. With thousands of genetic variants within the human genome and an equally large number of potential environmental factors that may affect lung function, the risk of false discovery is high when seeking evidence of gene-environment interactions. To minimize this risk, our study reduces uncertainty by testing only replicated genetic modifiers of lung function and the replicated environmental factor of ambient temperature.

We hypothesize the association of temperature and lung function in CF differs by genotype in selected modifier genes. In this study we tested for gene-environment interactions in several replicated loci of CF lung disease including two loci (chromosomes 11p13 and 20q13.2) identified through genome-wide association and linkage studies by the North American CF Modifier Consortium,[52] and polymorphisms in the mannose binding lectin (*MBL2*) gene supported through meta-analysis,[53] in the transforming growth factor beta-1 (*TGFB1*) gene supported through multiple studies,[54-58] and in interferon-related developmental regulator 1 (*IFRD1*) gene identified via genome-wide association and replicated in an independent population.[59] Within our CF study population we also tested for interactions in several modifiers of lung function replicated in the general population in large (n=7691 to 20,890) population-based genome-wide

association studies of lung function (FEV_1 and FEV_1/FVC ratio), including hedgehog interacting protein (*HHIP*), advanced glycosylation end product-specific receptor (*AGER*), C-terminal domain containing glutathione S-transferase (*GSTCD*), and 5-hydroxytryptamine receptor 4 (*HTR4*).[60-63] Together these analyses may provide information on why lung function is highly variable in CF. The above analyses were carried out in the CF Twin-Sibling Study population as this represents the largest sample of CF subjects with genetic data recruited over a wide temperature range to date.

Section 3.3. Methods

Study Population: This study was approved by the Johns Hopkins University Institutional Review Board (NA_00035659, NA_00019677). Subjects were recruited from CF centers based on having a twin or sibling also with CF (Table 3.1: n=2086 in 1018 families).[28] Data were collected between 10/27/00 and 3/31/13 with data supplementation from the U.S. CF Foundation Patient Registry through 12/31/11. Subjects were excluded if lung function, residential postal/zip code data, health insurance coverage data were not available, or if actively smoking.

Table 3.1. Study Population

Mean ± S.D.	**Study Population** (n = 1660)	**Excluded Subjects** (n = 426)	***P* value**
Sex (% male)	51.8%	56.6%	0.08
Race (% any non-white)	7.9% (n = 1637)	8.9% (n = 417)	0.53
Ethnicity (% Hispanic)	5.1% (n = 1659)	5.9% (n = 424)	0.56
Age at Diagnosis (years)	2.3 ± 5.5 [0.0, 52.0]	1.8 ± 4.3 [0.0, 33.0] (n = 424)	**0.043**
***CFTR* genotype** (% *F508del* homozygotes)	48.7%	45.7% (n = 418)	0.28
Pancreatic Insufficiency (%)	86.8% (n = 1650)	80.1%	**<0.001**
FEV_1 (CF-specific percentile)	69.5 ± 27.0 [0, 100]	67.7 ± 28.9 [0, 100] (n = 234)	0.36
Age (years at the time of FEV_1 measurement)	19.2 ± 9.6 [6.0, 63.9]	16.9 ± 9.0 [6.0, 48.4] (n = 234)	**0.001**
Insurance Status (% Public)	31.6%	35.9% (n = 273)	0.16
Mean Annual Temperature (°F)	54.2 ± 7.0 [38.5, 76.5]	55.5 ± 7.6 [42.6, 76.5] (n = 241)	**0.007**

Genetic Data: Genotyping has been previously performed for a sample of the Twin-Sibling Study population using the Illumina 610-Quad Array®, 660W-Quad Array®, and the Omni5 Array® under stringent quality control.[52] Two of the known genetic modifiers of CF lung disease have been identified using these arrays, specifically SNPs rs12793173 and rs4811626.[52] For the other three known CF loci, polymorphisms for a subset of the Twin-Sibling Study have been previously typed. These include codons 52, 54, and 57 in exon 1 (D, B, and C variant alleles, respectively) and promoter polymorphisms at positions -550 (H/L variant) and -221 (X/Y variant) for *MBL2*, -509

promoter site and codon 10 in *TGFB1*, and the rs7817 polymorphism in *IFRD1*.[33,59,87] The known genetic modifiers of lung function in general population to be examined include *HHIP*, *AGER*, *GSTCD*, and *HTR4*.[60-63] Actual genotype data was not available for all SNPs (Table 3.2), so imputation was performed using MaCH/Minimac software (http://www.sph.umich.edu/csg/abecasis/MACH/index.html). Phase I, Version 3 haplotype data from 1000 Genomes project (ftp://ftp-trace.ncbi.nih.gov/1000genomes /ftp/release/20110521/) was used as the reference. Samples were imputed separately by genotyping platform. Genotyped SNPs with a low minor allele frequency and low call rate were excluded prior to imputation. Imputed SNPs with a MaCH quality score r^2<0.30 were excluded from the analysis. For the purposes of analyses, SNPs were coded in an additive manner (0=homozygous for the major allele, 1=heterozygous for the minor allele, 2=homozygous for the minor allele). An additive model was selected as being more agnostic than dominant or recessive models. In addition, a composite variable was generated to estimate MBL "sufficiency" as previously described by our group.[87] Specifically, diplotypes of amino-acid substitutions associated with the minor alleles in codons 52, 54, and 57 (B,C, and D) and specific alleles at promoter site -221 may lead to lower levels of circulating MBL ("insufficient").

Table 3.2. Imputation and Minor Allele Frequency

Gene	SNP	Major/Minor Allele	Total Genotypes (n)	Imputed Genotypes (%)	Study Population Minor Allele Frequency (%)	dbSNP Minor Allele Frequency (%) [n = 1094 genotypes]	*P* value
Chromosome 11 Locus	rs12793173	T/C	1594	2.8	22.11	16.99	<0.001
Chromosome 20 Locus	rs4811626	C/T	1594	0.0	44.07	34.48	<0.001
IFRD1	rs7817	C/T	1594	2.9	51.70	39.30	<0.001
MBL2 **(Promoter -550)**	rs11003125	G/C	1619	31.0	36.90	30.61	<0.001
MBL2 **(Promoter -221)**	rs7096206	C/G	1619	31.0	20.40	19.55	0.44
MBL2 **(codon 52)**	rs5030737	C/T	1619	28.4	6.69	2.72	<0.001
MBL2 **(codon 54)**	rs1800450	G/A	1660	0.0	14.40	12.20	0.02
MBL2 **(codon 57)**	rs1800451	G/A	1619	31.0	2.79	8.13	<0.001
TGFB1 **(Promoter -509)**	rs1800469	C/T	1598	33.3	31.90	36.80	<0.001
TGFB1 **(codon 10)**	rs1800470	T/C	1538	41.9	41.28	45.47	0.002
HHIP **(SNP 1)**	rs12504628	T/C	1340	100.0	40.96	30.39	<0.001
HHIP **(SNP 2)**	rs13147758	A/G	1340	100.0	40.56	28.85	<0.001
HHIP **(SNP 3)**	rs1980057	C/T	1340	96.7	40.48	28.61	<0.001
GSTCD **(SNP 1)**	rs10516526	A/G	1594	2.8	6.12	3.17	<0.001
GSTCD **(SNP 2)**	rs11097901	C/T	1340	100.0	6.16	3.61	<0.001
HTR **(SNP 1)**	rs11168048	T/C	1340	96.7	41.94	43.01	0.45
HTR **(SNP 2)**	rs3995090	A/C	1594	2.8	41.98	44.37	0.08
AGER	rs2070600	G/A	1594	0.0	4.77	7.25	<0.001

Environmental Data: U.S. subjects were mapped to the center of their most recent known residential zip code (finest resolution available) to derive measures for mean

annual ambient temperature (°F). Source data for mapping was 1981-2010 mean annual ambient temperatures (The Climate Source; Corvallis, OR; 400 meter resolution).

Outcome Variables: Raw FEV_1 (liters) measurements were converted into CF-specific FEV_1 percentiles excluding measurements obtained before 6 years of age and after lung transplantation.[50] Lung function was defined as the best percentile in the most recent year of data.

Data Analysis: Regressions clustered by family (CFTSS) to account for familial relationships, chi square tests, and student's t-tests were performed using Stata 11 (StataCorp LP; College Station, TX). The interaction between each SNP and temperature was assessed through clustered linear regression with specific SNPs and temperature (and an interactive term derived by multiplying them) as the independent variables and lung function as the dependent variable; these regressions were also adjusted for age at the time of lung function and health insurance status based on previously published models.[44] Although this is single hypothesis testing with relatively few genetic modifiers being tested for interactions, Bonferroni correction for multiple tests was performed to provide the most conservative estimates of association.

Section 3.4. Results

Demographics: A total of 2086 individuals have been consented into the CF Twin and Sibling Study, of which 1660 (79.6%) met study inclusion criteria. Included subjects were 51.8% male and 7.9% were self-reported as non-white (Table 3.1). In terms of *CFTR* genotype, 48.7% were homozygous for the most common allele F508del similar to the general CF population. The mean FEV_1 specific percentile was 69.5±27.0, which is

higher than in the general CF reference population.[50] Included subjects had an older age of diagnosis (2.3 years vs. 1.8 years; p=0.043) and older age at the time of pulmonary function testing (19.2 years vs. 16.9 years; p=0.001) than excluded subjects. Included subjects also were more likely to have pancreatic insufficiency (86.8% vs. 80.1%; p<0.001) and reside in slightly cooler climes on average (54.2 vs. 55.5°F; p=0.007) than excluded subjects. There were no differences in frequencies or means between included and excluded subjects in terms of sex, race/ethnicity, *CFTR* genotype, or CF-specific FEV_1.

SNP Data: Owing to the different sources of genotyping, the 18 SNPs of interest had differing frequencies for imputed data, ranging from some SNPs being fully imputed to others being completely genotyped (Table 3.2). The minor allele frequencies statistically differed from reported frequencies in dbSNP for all but 3 SNPs, the two in the *HTR* gene and the *MBL2* -221 promoter site. For analysis, a composite variable for MBL2 sufficiency status was derived from 4 *MBL2* SNPs (-221 promoter site, and codons 52, 54, and 57) with 247 subjects described as "insufficient" (15.3%) and 1372 described as "sufficient" (84.7%). Although gene-environment interactions were tested for in 14 SNPs and the composite MBL2 status variable, it is recognized that these tests may not be independent tests owing to linkage within genes. In our study there were 5 genes/loci of interest which contained 2 or more SNPs of interest; correlation between these SNPs was assessed (Table 3.3). Conservatively, we assumed only SNPs with very high correlation would be considered dependent tests for statistical purposes, which would include the SNPs on the *HHIP, GSTCD,* and *HTR* genes. Thus, the subsequent interaction testing was Bonferroni corrected for performing 11 tests (i.e., one each for

HHIP, GSTCD, and *HTR* gene SNPs, one for the composite *MBL2* variable, and one each for the remaining 7 SNPs).

Table 3.3. Correlation (r^2) of SNPs within Genes

R-squared	*MBL2* Sufficiency Status	*TGFB1* (10)	*HHIP* (SNP2)	*HHIP* (SNP3)	*GSTCD* (SNP2)	*HTR* (SNP2)
***MBL2* (-550)**	0.046 (n=1619)	-	-	-	-	-
***TGFB1* (-509)**	-	0.677 (n=1538)	-	-	-	-
***HHIP* (SNP1)**	-	-	0.953 (n=1340)	0.939 (n=1340)	-	-
***HHIP* (SNP2)**	-	-	-	0.985 (n=1340)	-	-
***GSTCD* (SNP1)**	-	-	-	-	0.990 (n=1340)	-
***HTR* (SNP1)**	-	-	-	-	-	0.987 (n=1340)

Gene-Environment Interactions: Clustered linear regressions were performed using each SNP as an independent variable and lung function as the dependent variable (Table 3.4). In univariate analysis, two SNPs were associated with lung function (*TGFB1* -509 promoter site and *GSTCD* SNP rs11097901). These SNPs were also associated with lung function in multivariate regression after adjustment for age at the time of lung function testing, mean annual ambient temperature, and insurance status. However, assuming a Bonferroni corrected p-value of 0.0045 (0.05/11 independent tests), none of these SNPs remained associated with lung function after correction for multiple tests. To test for interactions, an interaction term (temperature-SNP) was added to the multivariate regression models for each SNP. Although 1 SNP and its interaction term was associated with lung function (*IFRD1*), this did not remain significant after correction for multiple testing.

Table 3.4. Linear Regressions for Change in Lung Function Percentile by SNP

Coefficient (Change in Lung Function Percentile) [Range][1]	n	Coefficient for SNP by Univariate Regression[2]	Coefficient for SNP by Multivariate Regression[3]	Coefficient for SNP by Multivariate Regression with Interaction[4]	Coefficient for Interaction Term
Chromosome 11 Locus	1594	-0.73 [-3.22, 1.77] (p = 0.57)	-0.81 [-3.26, 1.64] (p = 0.52)	1.25 [-20.09, 22.59] (p = 0.91)	-0.04 [-0.43, 0.36] (p = 0.85)
Chromosome 20 Locus	1594	-0.11 [-2.15, 1.93] (p = 0.92)	-0.37 [-2.39, 1.64] (p = 0.72)	-11.62 [-27.50, 4.25] (p = 0.15)	0.21 [-0.09, 0.50] (p = 0.17)
IFRD1	1594	0.04 [-2.08, 2.17] (p = 0.97)	-0.26 [-2.32, 1.80] (p = 0.81)	**-22.45** **[-40.44, -4.46]** **(p = 0.015)**	**0.41** **[0.07, 0.74]** **(p = 0.017)**
***MBL2* (Promoter -550)**	1619	-1.37 [-3.51, 0.76] (p = 0.21)	-1.58 [-3.64, 0.49] (p = 0.14)	-2.37 [-20.82, 16.07] (p = 0.80)	0.01 [-0.32, 0.35] (p = 0.93)
***MBL2* (Sufficiency Status)[5]**	1619	2.13 [-1.55, 5.82] (p = 0.28)	2.56 [-1.11, 6.25] (p = 0.17)	4.37 [-25.06, 33.80] (p = 0.77)	-0.03 [-0.57, 0.51] (p = 0.90)
***TGFB1* (Promoter -509)**	1598	**-3.10** **[-5.30, -0.91]** **(p = 0.006)**	**-2.92** **[-5.04, -0.79]** **(p = 0.007)**	11.03 [-5.97, 28.03] (p = 0.20)	-0.26 [-0.57, 0.05] (p = 0.10)
***TGFB1* (codon 10)**	1538	-2.08 [-4.21, 0.06] (p = 0.06)	-2.04 [-4.12, 0.05] (p = 0.06)	10.72 [-6.75, 28.20] (p = 0.23)	-0.24 [-0.55, 0.08] (p = 0.15)
***HHIP* (SNP 1)**	1340	1.31 [-1.02, 3.63] (p = 0.27)	1.32 [-0.97, 3.62] (p = 0.26)	9.45 [-7.69, 26.60] (p = 0.28)	-0.15 [-0.46, 0.16] (p = 0.34)
***HHIP* (SNP 2)**	1340	1.32 [-0.96, 3.61] (p = 0.26)	1.42 [-0.83, 3.67] (p = 0.22)	10.37 [-6.61, 27.35] (p = 0.23)	-0.16 [-0.47, 0.14] (p = 0.29)
***HHIP* (SNP 3)**	1340	1.51 [-0.77, 3.78] (p = 0.19)	1.60 [-0.64, 3.83] (p = 0.16)	10.60 [-6.33, 27.53] (p = 0.22)	-0.17 [-0.47, 0.14] (p = 0.29)
***GSTCD* (SNP 1)**	1594	3.04 [-1.23, 7.32] (p = 0.16)	2.79 [-1.40, 6.97] (p = 0.19)	-21.96 [-57.11, 13.22] (p = 0.22)	0.46 [-0.20, 1.11] (p = 0.17)
***GSTCD* (SNP 2)**	1340	**6.23** **[1.75, 10.71]** **(p = 0.007)**	**5.74** **[1.25, 10.23]** **(p = 0.012)**	-25.98 [-62.60, 10.64] (p = 0.16)	0.58 [-0.09, 1.26] (p = 0.09)
***HTR* (SNP 1)**	1340	0.13 [-2.21, 2.46] (p = 0.92)	0.17 [-2.16, 2.51] (p = 0.89)	-12.92 [-31.44, 5.59] (p = 0.17)	0.24 [-0.09, 0.57] (p = 0.16)
***HTR* (SNP 2)**	1594	-0.14 [-2.19, 1.92] (p = 0.90)	0.04 [-2.02, 2.11] (p = 0.97)	-6.97 [-23.05, 9.10] (p = 0.40)	0.13 [-0.16, 0.42] (p = 0.38)
AGER	1594	-2.29 [-7.54, 2.97] (p = 0.39)	-2.46 [-7.53, 2.62] (p = 0.34)	-15.16 [-61.19, 30.87] (p = 0.52)	0.23 [-0.61, 1.08] (p = 0.59)

[1]Based on the data presented in Table 3.3, it is assumed that 11 independent tests are being performed, thus the Bonferroni corrected p-value is 0.05/11 or 0.0045. None of the SNP co-efficient or interaction term p values meet this threshold.
[2]All regressions include lung function (FEV1) as the dependent variable and genotype for a SNP coded by the number of minor alleles as the independent variable, and are clustered by family.
[3]Multivariate regression includes adjustment for age at the time of lung function measurement in years, mean annual ambient temperature in degrees Fahrenheit, and insurance status coded as 0 for private insurance and 1 for public insurance.
[4]The interaction term is derived by multiplying the SNP variable and the value for temperature.
5MBL Sufficiency Status is a composite variable with MBL sufficiency or insufficiency defined on the basis of MBL2 diplotypes incorporating data from *MBL2* SNPs in the -221 promoter site, and codons 52, 54, and 57.

Section 3.5. Discussion

Lung disease in cystic fibrosis is a complex phenotype, influenced by mutations in the disease-causing *CFTR* gene itself, polymorphisms or mutations in other modifier genes, and environmental/stochastic factors.[29] We have previously described mean annual ambient temperature as an environmental factor associated with lung function in CF.[44] Although a portion of this association is explained by the geographic distribution of common respiratory pathogens in CF, the remaining mechanism(s) are unknown. We hypothesized that the remaining unexplained portion of the association between temperature and CF lung function might be secondary to gene-environment interactions between temperature and known modifier loci/genes in CF. However, regression-based testing did not identify any significant gene-environment interactions for any of the 14 tested SNPs or the composite MBL2 status variable.

The findings may be truly negative as the association between temperature and lung function in CF may not be mediated through the tested genetic modifiers, but rather the association might be mediated by other untested modifier genes or other environmental factors. Alternatively, there is a relationship, but we are underpowered to detect gene-environment interactions with a study population of 1660 subjects with genotypes with a minor allele frequency of less than 20%. Assuming a modest main genetic effect of 2.0 percentile points per minor allele, a main environmental effect of 0.3 percentile points per temperature degree based on previously published data, and a relatively robust gene-environment interaction effect of 0.5 percentile points (a decrease of an additional 0.5 percentile points of lung function per degree of temperature with an at-risk genotype), a post-hoc power calculation using Quanto[88] suggests that we had

>80% power to detect interactions with SNPs with a minor allele frequency >0.2 and >60% power for SNPs with a minor allele frequency >0.1. It is also possible that our Bonferroni correction for multiple testing is too stringent.

Perhaps the biggest study limitation was misclassification bias when assigned mean annual ambient temperatures. Although subjects may reside in multiple locations over a lifetime, lung function from the most recent year of data and each subject's last known location were used to minimize this uncertainty. Furthermore, in a prior study we found that 79.5% of the CF Twin-Sibling Study subjects live within the state in which they were born with another 8% living within an adjacent state (n=1313).[44] Other potential limitations include the use of imputed SNP data, although the correlation between observed and imputed alleles with a similar 660K platform is 0.93, suggesting that imputing a marker for 1000 subjects would provide similar data to genotyping that marker for 930 subjects.[64,65] It is also possible that the association between temperature and lung function is mediated by one of the studied SNPs through a non-additive model, which was not employed in this study as we felt an additive model to be more agnostic than dominant or recessive models.

Although we did not identify any gene-environment interactions in our study, we hope that our approach highlights the importance of testing for gene-environment interactions with both replicated genetic and environmental modifiers to minimize false positive findings. It is still remains important to identify gene-environment interactions in CF as well as elucidate the relationship between temperature and lung function in CF to identify susceptible individuals who could benefit from specific interventions to improve longevity.

CHAPTER 4. WARMER TEMPERATURES ARE ASSOCIATED WITH LOWER LUNG FUNCTION IN HEALTHY INDIVIDUALS

Section 4.1. Abstract

Background: Climate change may have many effects on human health, including lung function. Previous work has shown an association between warmer ambient temperature and lower lung function (FEV_1) in individuals with cystic fibrosis. We hypothesized that a similar relationship may exist within the general population.

Methods: Spirometry (FEV_1) data from two National Health and Nutrition Examination Survey cohorts representative of the U.S. non-institutionalized population (NHANES III: 1988-94; NHANES continuous: 2007-10) were merged with mean ambient annual temperature data and associations were assessed for using survey-weighted multivariate regression.

Results: After adjusting for potential demographic confounding factors, warmer temperatures were associated with lower lung function in both the NHANES III population *(p*=0.016; n=10,628) and the NHANES 2007-2010 population (*p*=0.002; n=9024). The effect was modest in both populations with a 0.7% and 1.0% predicted FEV_1 decrease for every 10°F increase in mean annual ambient temperature in the NHANES III and NHANES 2007-2010 populations, respectively.

Conclusions: Temperature influences lung function in the general population. While the influence may be modest in the general population, the effect size of the association may be greater in populations with respiratory disease, such as cystic fibrosis,

which may have ramifications for other obstructive respiratory disorders, such as asthma and COPD.

Section 4.2. Introduction

The primary function of the respiratory tract is gas exchange. A wide variety of tests have been developed to assess pulmonary function including spirometry which provides quantitative estimates of lung volumes and air flow obstruction. As spirometry is a commonly employed test in both primary care and subspecialty offices as a measure of lung disease (e.g., asthma, COPD, cystic fibrosis, pulmonary fibrosis, etc.), it is important to understand the factors that influence it.

Studies of monozygous twins who share 100% of their genes and dizygous twins who share 50% of their genes have produced differing estimates of heritability with some studies demonstrating more limited genetic effects on lung function,[89,90] and others suggesting that genetic factors may account for over half of the variation in lung function.[91,92] Taken together these studies imply that lung function is complex trait with both genetic and environmental factors contributing to its variation. A number of environmental factors that impact lung function on an individual level have been described, such indoor combustion,[93] environmental tobacco smoke,[94] and proximity to traffic.[95-97] Fewer environmental factors that impact lung function in large numbers of individuals simultaneously over a wider geographical area have been described, but include air pollution.[96,98] Another large-scale geographic factor which may impact lung function, and respiratory health overall, is climate.

In the context of climate change, it is important to understand the role of temperature on disease.[99] Ambient temperature has been demonstrated to be associated with the prevalence of both infectious and non-infectious diseases,[45] particularly vector-borne infectious diseases. The acute effects of temperature on respiratory diseases has been perhaps best studied in COPD with increased mortality and morbidity with both hot and cold temperature extremes.[100] Acute temperature changes may also have an effect on respiratory symptoms in asthma.[101,102] Potential long-term effects of ambient temperature on lung function are less well studied. However, we have previously demonstrated a novel association between lower lung function (FEV_1) in individuals with cystic fibrosis and warmer annual ambient temperatures with replication in two national cystic fibrosis registries in the United States and Australia.[44]

We hypothesize that mean annual ambient temperature is associated with variation in lung function in healthy individuals without cystic fibrosis or other lung diseases. To test this hypothesis, we assessed cross-sectional lung function (FEV_1) in healthy non-smokers in the United States using two separate cohorts from the National Health and Nutrition Examination Survey (NHANES), which collected both geographic and spirometry data.

Section 4.3. Methods

Study Population: De-identified data were utilized from the National Health and Nutrition Examination Survey (NHANES) conducted by the National Center for Health Statistics (NCHS). NHANES is a series of surveys that collect questionnaire and medical testing data on a sample of the non-institutionalized U.S. population. All surveys where

spirometry was routinely performed were used, including the periods 1988-1994 (NHANES III), 2007-2008, and 2009-2010. The NHANES 2007-2008 and 2009-2010 data were combined for analysis as they utilize similar survey weights. The primary population was the NHANES III study population as a "representative" sample of the U.S. general population. The NHANES 2007-08 and 2009-10 data were used to replicate the analysis. To obtain a sample of participants in good respiratory health, participants were excluded if they had any mention of asthma or any reported active tobacco product use. Of note, specific questions for COPD were not asked of the study populations.

Demographic Data: Raw data for sex, age, race/ethnicity, annual household income, and insurance status were all obtained from NHANES publically available datafiles. Any participant reporting any non-white race or Hispanic ethnicity was coded as non-white for the purposes of regression modeling. Participants who had insurance were coded as either having no insurance, public insurance, or private insurance with any mention of private insurance being coded as private for the purposes of regression modeling.

Lung Function Data: Raw forced expiratory volume in 1 second (FEV_1) in milliliters was converted to percent predicted using NHANES reference equations.[46,103] Participants without valid FEV_1 data were excluded.

Temperature Data: Temperature source data consisted of a 30 year average of temperature (1980-2010; The Climate Source, Inc.; Corvallis, OR) to minimize year-to-year variation. Participants were geocoded for annual ambient temperature to the

population centroid of their last known residential postal zip code using ESRI ArcGIS (ESRI; Redlands, CA).[44]

Statistical Analyses: Student's *t* tests and chi-square tests were used to compare demographic factors between included and excluded participants. Linear regression was performed to determine if temperature was associated with lung function using separate models for the NHANES III data and the NHANES 2007-2010 data. The final models adjusted for demographic factors that were associated with lung function in univariate regression, including sex, race/ethnicity, age, the log value of household income, and insurance status if these factors were associated with lung function in univariate models. Although participants are selected to represent the U.S. population at all ages, NHANES routinely oversamples persons aged 60 years and older, African-Americans, and Hispanics, thus all regressions were survey weighted per recommended NHANES analytic practices.[104] Random effects meta-analysis was used to provide a combined estimate of effect size from both NHANES datasets.[105] All statistical analyses were conducted using STATA IC 11 (College Station, TX) at the NCHS RDC facility (Hyattsville, MD) owing to the use of restricted data (zip code). A p value of <0.05 was considered statistically significant.

Section 4.4. Results

Demographics: The NHANES III sample consisted of 25,733 participants over 6 years of age who could have had spirometry measurements, of which 4446 did not have spirometry performed for administrative reasons or medical exclusions. We excluded an additional 6432 participants, including 1315 participants without zip codes that could not

be geocoded for mean annual ambient temperature, 727 participants with asthma, and 4390 participants with active tobacco use, to yield a NHANES III study population of 14,855 participants whose data were used for analyses. The NHANES 2007-2010 sample consisted of 17,244 participants over 6 years of age who could have had spirometry measurements, of which 4972 did not have spirometry performed for administrative reasons or medical exclusions. We excluded an additional 3248 participants, including 3 participants without zip codes that could not be geocoded for mean annual ambient temperature, 1034 participants with asthma, and 2211 participants with active tobacco use, to yield a NHANES 2007-2010 study population of 9024 participants whose data were used for analyses.

Survey weighted means for selected demographic variables, which account for NHANES oversampling of certain demographic groups, are reported in Table 4.1. Comparisons between excluded and included populations are reported in Table 4.2. The differences between the included and excluded populations largely related to the exclusion of participants with active tobacco use, which was more common among males, Blacks and Whites, older participants, and participants with lower reported incomes or respiratory conditions in both the NHANES III and NHANES 2007-2010 samples.

Table 4.1. Survey Weighted Means for Study Populations

Mean	NHANES III Data	NHANES 2007-2008 and 2009-2010 Data
	n = 14,855	n = 9024
Sex (% female)	54.7	53.0
Race/Ethnicity (%)		
Black	11.5	10.4
Hispanic	6.6	16.2
White	81.8	73.4
Age (years)	37.4	37.6
Income ($'000s)	31.5 (n=13,526)	60.1 (n=8668)
Insurance (%)		
None	11.1	15.3
Private	66.1	68.3
Public	22.8 (n=14,227)	16.4 (n=9007)
Temperature (°F)	57.2	56.7
FEV_1 (% Predicted)	98.6	99.0

Table 4.2. Included and Excluded Populations

Mean ± S.D.	NHANES III Data		NHANES 2007-2008 and 2009-2010 Data	
	Included Participants (n = 14,855)	Excluded Participants (n = 10,878)	Included Participants (n = 9024)	Excluded Participants (n = 8220)
Sex (% female)	55.6	48.4	52.4	48.3
Race/Ethnicity (%)				
Black	28.0	30.2	18.9	22.5
Hispanic	31.7	23.4	35.8	27.2
White	40.3	46.4	45.3	50.4
Age (years)	37.1 ± 23.5	42.5 ± 24.2	35.1 ± 22.0	43.1 ± 23.6
Income ($'000s)	25.4 ± 15.6 (n=13,526)	23.3 ± 15.3 (n=9554)	50.2 ± 32.1 (n=8668)	40.3 ± 30.0 (n=7797)
Insurance (%)				
None	17.8	16.6	20.0	23.1
Private	50.6	45.6	54.7	44.3
Public	31.6 (n=14,227)	37.9 (n=10,304)	25.3 (n=9007)	32.6 (n=8208)
Asthma (% yes)	0.0	8.5	0.0	18.4
Tobacco Use (% yes)	0.0	52.4	0.0	37.0
Temperature (°F)	60.1 ± 7.7	59.5 ± 7.6 (n=9943)	58.3 ± 8.4	57.8 ± 8.3 (n=8215)
FEV_1 (% Predicted)	99.4 ± 15.7	92.7 ± 18.0 (n=5845)	100.0 ± 14.2	93.1 ± 16.4 (n=3248)

Potential Confounders: Survey-weighted univariate modeling of the association of lung function and various demographic factors was performed to identify potential confounding factors (Table 4.3). The two study populations demonstrated a similarity in

the effect of confounders such as race and age while higher reported incomes and private insurance were associated with higher lung function only in the NHANES III population, whereas female sex was associated with higher lung function only in the NHANES 2007-2010 population. For the multivariate regression analyses that follow significant factors were included as covariates.

Table 4.3. Survey Weighted Univariate Regressions

Coefficient (change in FEV_1 % predicted) [95% C.I.]	**NHANES III Data**		**NHANES 2007-2008 and 2009-2010 Data**	
	n = 14,672	Coefficient *P* value	n = 9024	Coefficient *P* value
Sex (male=0; female=1)	-0.19 (-0.84, 0.45)	0.55	1.34 [0.80, 1.89]	<0.001
Race/Ethnicity (white=0; nonwhite=1)	2.41 [1.52, 3.30]	<0.001	2.26 [1.44, 3.07]	<0.001
Age (per year)	-0.12 [-0.14, -0.11]	<0.001	-0.14 [-0.16, -0.12]	<0.001
Log Income (per Log $)	0.90 [0.01, 1.79] (n=13,354)	0.047	-0.72 [-2.01, 0.57] (n=8668)	0.26
Insurance (private=0; public/none=1)	-2.37 [-3.08, -1.65] (n=14,048)	<0.001	0.19 [-0.73, 1.12] (n=9007)	0.67

Relationship of Lung Function with Temperature: In survey weighted univariate modelling, higher temperatures were not associated with lower lung function in the NHANES III population *(p*=0.19) or the NHANES 2007-2010 population (*p*=0.17). In survey weighted multivariate regression incorporating the covariates identified through univariate (Table 4.4), higher temperatures were associated with lower lung function in both the NHANES III population *(p*=0.016) and the NHANES 2007-2010 population (*p*=0.002). The effect was modest in both populations with a 0.7 and 1.0 percentile point FEV_1 decrease for every 10°F increase in mean annual ambient temperature in the NHANES III and NHANES 2007-2010 populations, respectively. Meta-analysis

incorporating both study results yielded an estimated effect size of 0.9 percentile point decrease for every 10°F increase (95% C.I.: 0.5%, 1.3%; $p<0.001$).

Table 4.4. Survey Weighted Multivariate Regressions Adjusted for Potential Confounders for Study Population

Coefficient (change in FEV_1 % predicted) [95% C.I.]	**NHANES III Data**		**NHANES 2007-2008 and 2009-2010 Data**	
	n = 10,628	Coefficient *P* value	n = 9024	Coefficient *P* value
Sex (male=0; male=1)	Not adjusted for sex as this was non-significant in univariate model		1.48 [0.97, 1.99]	<0.001
Race/Ethnicity (white=1; nonwhite=1)	2.20 [1.35, 3.04]	<0.001	1.92 [1.12, 2.72]	<0.001
Age (per 10 years)	-1.04 [-1.29, -0.80]	<0.001	-1.38 [-1.59, -1.17]	<0.001
Log Income (per Log $)	1.69 [0.53, 2.85]	0.005	Not adjusted for log income as this was non-significant in univariate model	
Insurance (private=0; public=1)	-0.70 [-1.93, 0.53]	0.26	Not adjusted for log income as this was non-significant in univariate model	
Temperature (per 10°F)	-0.74 [-1.34, -0.15]	0.016	-1.01 [-1.62, -0.40]	0.002

Section 4.5. Discussion

In our analyses of cross-sectional NHANES data from two time points, 1988-1994 and 2007-2010, we observed an association between warmer temperatures and lower lung function in study populations representative of the U.S. non-institutionalized population. The effect was modest in both populations with a 0.7 and 1.0 percentile point FEV_1 decrease for every 10°F increase in mean annual ambient temperature in the NHANES III and NHANES 2007-2010 populations, respectively. We had previously reported an association between lower lung function and warmer ambient temperatures in several independent samples of individuals with cystic fibrosis in the United States and Australia.[44] We previously estimated that a hypothetical 18 year old white male with cystic fibrosis (Height: 175 cm) with an FEV_1 of 73.5% predicted who lived in a cold climate would be expected to have an FEV_1 of 66.1% predicted if he resided in 30 degree (°F) warmer climate, which equates to approximately 2.5 percentile point decrease in

lung function being associated with a 10°F increase in mean annual ambient temperature, or greater than the 0.7 and 1.0 percentile point decreases seen in the general population. Taken together these results suggest that climatic temperatures may be associated with lung function in all individuals with a perhaps a more prominent effect in individuals with lung disease.

This difference may reflect that same factors operating in the general population that decrease lung function with warmer temperatures may be more potent on a background of cystic fibrosis, or conversely, in addition, to the factors operating in the general population, individuals with cystic fibrosis living in warmer climates may have other factors impacting their lung function. Of note, we did previously describe that certain respiratory pathogens (*Pseudomonas aeruginosa*) in cystic fibrosis do mediate a portion of the association (6-22% of the association) between lung function and temperature in cystic fibrosis. As *P. aeruginosa* does not typically cause respiratory illnesses in healthy individuals, this would be an example of a factor operating in cystic fibrosis and not the general population.

There are a number of limitations to this study. Specifically, there are limitations with regards to location of the NHANES participants. First, participants may reside in multiple locations over their lives, and it is unknown whether interactions between climate and age exist, specifically whether exposure at particular ages has greater consequences. Second, NHANES does sample from selected geographic regions throughout the United States, and the geographical distribution of participants may not be representative of the United States as a whole. Another limitation relates to our measurement of temperature. In order to minimize year-to-year variation in temperature,

we utilized a 30 year average. It may be that increased temperatures at certain times of the year (e.g., summer) may have greater effects on health.[106] Other limitations included potential confounding factors that we were unable to assess, such as physical activity, allergens, and air pollution.[100,107,108]

Our findings suggest that temperature may have a modest influence on lung function in the general population, although the mechanisms behind such an association remain unclear. While the influence may be modest in the general population, the effect size of the association may be greater in populations with respiratory disease, such as cystic fibrosis, which may have ramifications for other obstructive respiratory disorders, such as asthma and COPD. Although the average increase in temperature with climate change may be small, it may still lead to an increase in overall respiratory disease burden as some individuals may be more susceptible to increases in ambient temperatures.

CHAPTER 5. CONCLUSIONS

Section 5.1. Introduction

Respiratory diseases, both acute and chronic, are responsible for significant morbidity and mortality in both the developed and developing world. Pneumonia is the most common cause of death in infants and young children under the age of 5 worldwide.[109] The most common respiratory disease, asthma, will affect 100 million people worldwide by 2015 and COPD is predicted to become the third leading cause of death by 2030.[110] Although potent risk factors for the development and severity of respiratory diseases have been identified, such as tobacco smoke exposure and air pollution,[111] the development of most respiratory disorders remains a complex blend of genetic and environmental factors. Understanding the factors that impact respiratory diseases and their interactions (i.e., gene-gene, gene-environment, environment-environment) is an important first step in ameliorating the morbidity and mortality that accompanies them.

Another key issue in understanding the epidemiology of diseases is to also recognize changing patterns of disease so as to deploy resources more efficiently. For example, the prevalence of asthma among children in the United States doubled between 1980 and 1995, and continued to rise through 2009, although the reasons for this are not entirely clear.[112] Factors that lead to disease may also change over time. For example, ongoing public health campaigns and efforts to decrease smoking will undoubtedly lead to less smoking-related disease.[113] Another environmental factor which is anticipated to change in the upcoming decades is climate, specifically global warming.[114] In fact,

climate change may be the "biggest global health threat of the 21st century."[115] The World Health Organization has already estimated that climate change over the past 30 years is claiming over 150,000 lives annually.[45]

Temperature has been recognized a risk factor for worse outcomes in respiratory diseases.[99] Notably, acute respiratory events and exacerbations of existing respiratory disease are more likely to occur during times of temperature extremes. With heat waves and cold snaps, an increased frequency of exacerbations and/or deaths has been observed in asthma and COPD.[100-102] However, there are limited data on how climate affects long-term trajectories of common diseases. One potential disease model for studying the effects of climate on respiratory disease is cystic fibrosis (CF). We previously demonstrated that lower lung function is associated with warmer annual ambient temperatures in 3 independent populations of subjects with cystic fibrosis, including 2 populations in the United States and a third in Australia.[44]

The mean ambient temperature changes associated with anthropogenic global warming may be modest, 0.3-1.7°C under the lowest emission scenario and 2.6-4.8°C under the highest.[114] However, in a hypothetical subject with CF, there would be a 0.45 percent predicted decrease in lung function (FEV_1) for every 1°C increase in mean ambient temperature, which could potentially equate to an additional 1-2 years of survival based on current population-based estimates of lung function decline in CF using the highest emission scenario.[116] However, the mechanism for this association between climate (i.e., ambient temperature) and long-term lung function in CF is not known, nor is it known whether this phenomenon exists in other populations, including the general population without respiratory conditions.

CF serves as an excellent model for studying the effects of climate on respiratory diseases, owing to a lower likelihood of disease misclassification as it is well-characterized, highly penetrant Mendelian disorder with a causative gene that has been known for over 25 years.[1] Effects on lung function due to environmental exposures such as temperature may be magnified in a host predisposed to more rapid lung function decline. In addition, dense, longitudinal clinical information is collected routinely every three months for the entire lifetime of people with CF. Despite being a single gene disorder, CF lung disease remains a complex trait as patients with identical mutations may have dramatically different lung function,[2-4] implying the involvement of factors other than the *CFTR* gene itself. Family-based studies by our group and others comparing lung function between monozygous twins (who share 100% of their genes) and dizygous twins and siblings (who share 50% of their genes) suggest ~50% of variation in CF lung disease can be attributed to modifier genes and ~50% to environmental modifiers.[29,30] This allows for the possibility of studying environment-environment and/or gene-environment interactions with regards to climate.

A number of environmental factors that are associated with lung function (and could interact with climate) have been identified in the literature, including secondhand tobacco smoke,[33-37] household income,[39,70,71] occupation,[117] health insurance status,[39-42,44] parental education levels,[39] adherence to medical therapies,[118] exercise,[119-124] colonization with specific respiratory pathogens,[8-15] and air pollution.[38] Likewise, several replicated genetic loci have been identified that could interact with climate, which include two loci (chromosomes 11p13 and 20q13.2) identified through genome-wide association and linkage studies by the North American CF Modifier Consortium,[52] and polymorphisms in

the mannose binding lectin (*MBL2*) gene,[53] the transforming growth factor beta-1 (*TGFB1*) gene,[54-58] and in the interferon-related developmental regulator 1 (*IFRD1*) gene.[59] Overall, the presence of environmental and genetic modifiers in a readily diagnosed chronic lung disease characterized by short-term exacerbations and long-term decline makes CF a useful model to study the effects of climate on lung disease.

This study focused on three aspects of the association between temperature and lung function in cystic fibrosis. The first part (Chapter 2) focused on identifying environmental mediators of this association, the second part (Chapter 3) focused on identifying gene-environment interactions, and the third part (Chapter 4) focused on determining whether this association existed in non-CF populations. The mediation analyses in the first part of this study ascribed a portion of the association in cystic fibrosis to 3 respiratory pathogens. The second part of this study did not identify any gene-environment interactions between temperature and selected SNPs that are known modifiers of CF lung disease or lung function in the general population. The third part of the study confirmed the presence of this association within the general population though to a lesser degree; hence, there may be some relationship between lung function and temperature implicit in human biology or an unmeasured confounder with a geographic distribution similar to that of mean annual ambient temperature. The findings of this study have important implications for both individuals with cystic fibrosis with respect to their healthcare and the general population with respect to population health and climate change.

Section 5.2. Implications: Cystic Fibrosis

Taken together, the results from the first and third parts of this study (Chapters 2 and 4) could potentially explain 40-83% of the association between lower lung function and warmer temperatures in CF and identify some mechanisms that underlie this association, although it should be noted that the confidence intervals of the mediation analyses are wide. The mediation analyses in the first part of this study ascribe 12% or 43% of the association in cystic fibrosis to 3 respiratory pathogens, depending on the population of CF subjects examined. The third part of the study also confirms the presence of this association within the general population with effect sizes of 0.7% and 1.0% predicted FEV_1 decrease for every 10°F increase in mean annual ambient temperature in the 2 NHANES general population cohorts to an estimated 2.5% FEV_1 decrease in CF suggests that 28% or 40% of the association seen in CF can be ascribed to a mechanism also present in the general population, depending on the NHANES cohort examined.

Combining these estimates, 40-83% of the association between lung function and temperature may be ascribed to CF respiratory pathogens or mechanisms also seen in the general population, and 17-60% may be accounted for other as-of-yet unidentified factors particular to CF. This study did not identify any gene-environment interactions to explain the association (Chapter 3), but they still may exist as only a small selection of SNPs were tested and this study may be underpowered to detect small effect size interactions.

With regards to individuals with cystic fibrosis, the results from the first part of this study suggest that where patients reside impacts outcomes. Specifically, infection rates, either prevalence of organisms on respiratory cultures or incidence of initial infection, are not constant across the United States. The geographic stratification of

respiratory pathogens in CF is not unexpected given geographic distribution patterns of many infectious organisms. Recognizing differences in infection rates has ramifications for allocating resources in CF care and different approaches to care itself.

Different infection rates geographically should have an impact on benchmarking. In recent years, there has been increased interest in following health outcomes on a population level among policy-makers and healthcare payers among others.[125] The U.S. CF Foundation, which accredits CF care centers in the United States and promotes/spearheads best care practice efforts, has been comparing a variety of clinical outcomes on a population-basis between CF care centers, including lung function. The work in this study suggests that comparisons of centers may need to be weighted based on geographic location as there may be factors dictating respiratory pathogen prevalence and lung function that are macro-environmental factors outside of their control (i.e., ambient temperature). Likewise, payers of healthcare who reimburse under a capitated system whereby providers receive a lump annual sum to care for a population of patients may need to reimburse at a higher rate per patient for healthcare providers in warmer climates, perhaps owing to lower lung function and higher prevalence of certain respiratory pathogens.

In terms of CF management, the implications revolve around the earlier diagnosis of the presence of infectious organisms. Specifically, patients living in warmer climates may acquire certain respiratory pathogens at earlier ages, leading to a more rapid rate of lung function decline at an earlier age. Often with initial acquisition of certain respiratory pathogens in CF, including *P. aeruginosa*, eradication is possible,[126] thus surveillance via respiratory cultures may need to be more frequent to detect initial acquisition as early as

possible, or eradication and prevention may need to be more aggressive in clinics located in warmer climates. It should be noted that it also may be more difficult to eradicate or treat these respiratory pathogens in warmer climates, and perhaps longer treatment course of antimicrobial therapies could be warranted.

Our results may have implications for multi-center clinical research studies as well. For studies that are conducted across a diverse geographic area, as results may need to be adjusted for climatic variation. This may affect studies where the exposure of interest is also geographically distributed (i.e., exercise), otherwise false positive or negative results may be seen depending on whether studied exposure correlates or inversely correlates with temperature. Additionally, multi-center studies that are randomized by center may also be affected if there are unbalanced differences in temperature between sites as a whole. Strategies to address this include adjusting for temperature in analyses and/or randomizing by subject instead of by study site.

In addition to examining the infections as an environmental mediator between temperature and lung function, this study also tested for gene-environment interactions between temperature and previously replicated non-*CFTR* gene modifiers of CF lung disease and gene modifiers of lung function in the general population. As previously stated, family-based studies suggest ~50% of variation in CF lung disease can be attributed to modifier genes,[29,30] but significant loci within a GWAS study only explain an estimated 4-46% of the variability of CF lung function,[52] thus gene-environment interactions may explain the unexplained variation seen in CF lung disease.[127] In the second part of this study we tested for gene-environment interactions in several replicated loci of CF lung disease[52-59] and several modifiers of lung function replicated in the

general population.[60-63] However, regression-based testing did not identify any significant gene-environment interactions for any of the tested SNPs after correction for multiple testing. It should be noted that only a limited number of SNPs were tested, and a genome-wide study may identify an interaction(s). In addition, given that larger populations are frequently needed to identify gene-environment interactions compared to identifying simply loci of interest or an environmental exposure of interest,[128] more precise quantification of the environmental factor of temperature may be required to improve the power to detect an interaction.

Section 5.3. Implications: General Population

In the third part of this study we tested the hypothesis that mean annual ambient temperature was associated with variation in lung function in healthy individuals using cross-sectional lung function (FEV_1) in healthy non-smokers in the United States from two separate cohorts from the National Health and Nutrition Examination Survey (NHANES). We observed an association between warmer temperatures and lower lung function in study populations with an effect in both cohorts, specifically, a 0.7 and 1.0 percentile point FEV_1 decrease for every 10°F increase in mean annual ambient temperature in these 2 cohorts. Thus, assuming a 30°F difference in mean annual ambient temperature between colder northern regions in the continental United States and the warmer southern regions, individuals in warmer regions could have lung function that was 2.1-3.0 percentile points lower on average compared to those living in colder regions. This effect size was more modest compared to effect seen with reported tobacco use in both NHANES populations. In the NHANES III cohort, the percent predicted FEV_1 was 5.9 percentile points lower in those reporting tobacco use (n=4773) compared

to those who did not (n=15,927; *t*-test *p* value<0.001). Likewise, in the NHANES 2007-2010 cohort, the percent predicted FEV_1 was 6.4 percentile points lower in those reporting tobacco use (n=2427) compared to those who did not (n=9849; *t*-test *p* value<0.001).

With regards to the general population, the mechanism behind this association of lower lung function and warmer ambient temperatures remains unclear. It is possible that ambient temperature has a direct effect on lung function. For example, pulmonary function testing (FEV_1) in subjects with and without asthma can be affected by variation in temperature and humidity within calibrated environmental chambers.[129] Alternatively, the association may be mediated through another exposure such as exercise or indoor air pollution, which may be subject to regional differences, possibly owing to outside ambient temperature.[130,131] Additionally, it is possible that a confounding factor could lead to varying lung functions locally. For example, living in an urban area may lead to exposure to a higher mean annual ambient temperature compared to living in the same geographic spot if it were rural, and living in an urban area could also lead to greater exposure to air pollution, which adversely affects lung function.[96,97,132]

Another confounding factor may be the geographic distribution of respiratory pathogens; an increased incidence of respiratory infections in particular regions may lead to lower lung function for the population as whole even if only a limited number of people are affected. For example, limited data on the proportion of MRSA isolates among all *S. aureus* isolates suggests that there may be geographical trends in the distribution of the proportion of MRSA isolates from *S. aureus* isolates in the United States and Europe.[133,134] Lastly, given that the effect size of the association is greater in a lung

disease such as cystic fibrosis compared to the general population, it is possible that a greater effect size is also present with other lung diseases. While we attempted to exclude the most common chronic respiratory disease from the NHANES study populations in the third part of this study, it is possible that individuals with lung disease remain in the study population. If these individuals were geographically distributed such that certain respiratory diseases other than asthma (e.g., COPD) or environmental allergies were more prevalent in warmer climates, then this could have influenced our results.[135]

The implications on public health of our findings may be of increasing import in the future. As previously mentioned the mean ambient temperature changes associated with anthropogenic global warming may be 0.3-1.7°C under the lowest emission scenario and 2.6-4.8°C under the highest.[114] Thus, assuming the lowest emission scenario temperature change (0.3°C) and the lower lung function change with temperature (0.7%/10°F), then the average lung function decrease would be 0.04 percentile points. Conversely, assuming the highest emission scenario temperature change (4.8°C) and the larger lung function change with temperature (1.0%/10°F), then the average lung function decrease would be 0.86 percentile points. While these are very modest changes on the population level, it may still lead to an increase in overall respiratory disease burden as it is likely that some individuals may be more susceptible to increases in ambient temperatures. Specifically, this may still have ramifications for obstructive respiratory disorders, such as asthma and COPD,[100,136] where the effect size is likely greater. Certainly for cystic fibrosis, given our observations that warmer temperatures were associated with several respiratory pathogens, it is possible that global warming may lead

to a higher prevalence of specific pathogens among individuals with CF, and more rapid declines in lung function.

Section 5.4. Strengths and Limitations

This study has several strengths, including a well-characterized primary population (CF Twin-Sibling study) with respiratory, infectious, and genetic data that was used for the first two parts of the study, which focused on CF. The replication population for mediation with CF respiratory pathogens included the largest available database in term of numbers of participants with CF. The NHANES cohorts used for the third part of this study represent one of the largest databases of spirometric data, and has also been used as a standard reference population.[46,103] Spirometry is a commonly used measure of respiratory health, and lung function testing in patients with CF is generally robust with multiple tests performed annually with routine clinical care. The NHANES spirometric protocols are standardized with robust analysis algorithms.[137]

Our study also had several limitations, the most notable being the potential for misclassification bias by geographic location as both individuals with CF and NHANES participants move over the course of their lifetimes. Although subjects may reside in multiple locations over a lifetime, lung function from the most recent year of data and each subject's last known location were used to minimize this uncertainty. Furthermore, in a prior study we found that 79.5% of the CF Twin-Sibling Study subjects live within the state in which they were born with another 8% living within an adjacent state (n=1313).[44] However, it is unknown whether interactions between climate and age exist, specifically whether exposure (i.e., location) at particular ages has greater consequences.

It should also be noted that NHANES does sample from selected geographic regions throughout the United States, and the geographical distribution of participants may not be representative of the United States as a whole. Unfortunately, due to CDC restrictions on the use of data identifiers, we were unable to confirm the geographic distribution of NHANES participants.

Another limitation is the possibility of unmeasured confounders. For the first part of this study, which focused on respiratory pathogens, *P. aeruginosa* has been characterized as a highly diverse species,[82] and geographic variation in MRSA strains has been documented.[83] Thus it is possible that more virulent strains are geographically distributed in warmer climates by chance alone, and we did not account for strain differences. Likewise, the presence of non-tuberculous mycobacteria in CF respiratory cultures has been associated with unmeasured factor of vapor pressure, but not ambient temperature.[84] For all three portions of the study, other potential confounding factors that we were unable to assess include physical activity, allergens, and air pollution.[100,107,108] For example, there may be an interaction between air pollution and temperature,[100] which was unaccounted for in our model.

Measurement error, especially for temperature and infection, is another concern. While the measurement of pulmonary function may be fairly robust, our temperature measurement was a coarse measurement. Ambient temperature was estimated for individuals based on residential postal zip codes, but this may not be the actual temperature subjects were exposed to. Also, in order to minimize year-to-year variation in temperature for the purposes of study design, we utilized a 30 year average, but it may be that increased temperatures at certain times of the year (e.g., summer) may have

greater effects on health.[106] For first part of the study which incorporated infection data, it should be noted that these data were obtained from clinical specimens largely obtained through sputum samples and throat swabs; the latter, in particular, may not accurately represent lower airway flora, especially at younger ages.[18] However, poor quality of culture data would likely bias our results to the null, rather than leading to positive associations. Also, our use of retrospective culture data may under- or overestimate the prevalence of respiratory pathogens depending on reporting biases by CF care centers who may not report all culture data to the CF Foundation.

Another potential measurement error is the use of imputed SNP data in the second part of this study. However, the correlation between observed and imputed alleles with a similar 660K platform that was utilized is 0.93, suggesting that imputing a marker for 1000 subjects would provide similar data to genotyping that marker for 930 subjects.[64,65] Lastly, even though the measurements of lung function were collected using standardized protocols, it is critical to note that lung function measurements were cross-sectional owing to study design considerations of linking location and lung function temporally; however, a single value may not necessarily reflect the baseline lung function. This may be particularly the case for individuals with cystic fibrosis who have respiratory exacerbations with transiently lower lung function.[21]

Lastly, our models may not accurately represent the association between temperature and lung function. Adjusted clustered linear regression (generalized estimating equations) was utilized as the primary mode of analysis for all three parts of this study, and it is possible that the relationship between temperature and lung function was not linear. Additionally, our model for the first part of this study assumes that the

mediation variables of infection act upon the outcome variable of lung function, whereas it is also possible that more severe lung disease can lead to an increased propensity for infections. Our mediation analyses also had wide confidence intervals, so it is possible that we markedly underestimated or overestimated the impact of mediation with infectious pathogens on the association between temperature and lung function. It is also possible that individuals with CF may have polymorphisms in modifier genes affecting infection acquisition/colonization that are geographically distributed owing to founder effects or immigration patterns. Likewise, in the second part of the study, we assumed that gene-environment interactions occurred through additive models, rather than dominant or recessive. It is also possible that the association between temperature and lung function is mediated by one of the studied SNPs through dominant or recessive models, which was not employed in this study as we felt an additive model to be more agnostic than non-additive models.

Section 5.5. Future Directions

This study provides some evidence regarding potential mechanisms underlying a relationship between ambient temperature and lung function with several avenues of future research. The first avenue of research could focus on CF patients. Although 30-83% of the association between temperature and lung function can be attributed to respiratory pathogen distribution and general population effects, this still leaves work to be done identifying other factors. In addition, work could also focus on why respiratory pathogens are more prevalent in warmer climates among CF patients, such as whether this is due to an increased presence of these organisms in warmer climates (e.g., larger

reservoirs) or increased susceptibility to infection or colonization in warmer climates due to pathogen-specific or host-specific factors.

Although studies will be observational by nature as individuals cannot be randomized by climate, improved data regarding location over time could be very helpful. Specifically, data that is linked to an individual's location over a lifetime could parse out the effects of specific geographic regions and or age-climate interactions. Alternatively, studying outcomes in subgroups of individuals who have moved from a warmer climate to a colder climate or vice versa may provide more power to detect some of these effects, although there may be other confounding effects of environment. For example, many individuals may move upon entering adulthood to attend educational institutions some distance from their childhood residence, but the different environment of college may introduce additional effects.[29]

It is also possible that temperature may have an effect on CFTR function itself through the inhaled air temperature acting upon CFTR-bearing respiratory epithelial cells. There is some evidence that at least the most common mutation of *CFTR* (*F508del*) may have thermal instability.[138,139] Worse protein function at warmer temperatures could lead to worse disease outcomes as certainly functional classes are associated with differences in phenotype and mortality rates.[140] One potential *in vivo* measurement of CFTR function is sweat chloride.[141] Advances in technology may lead to wearable sweat chloride monitoring devices in the near term future.[142,143] Such devices would allow for real-time measurement of a biomarker for CFTR function under differing ambient temperature conditions, albeit in a different tissue than respiratory epithelium.

A second avenue of research could focus on the general population, including confirming this association in other regions of the world. Ideally, these regions would have similar healthcare systems across the population to be studied to minimize this source of variation. Given that the United States has an ethnically diverse population and lung function can be adjusted for race/ethnicity, studies may not need to be carried out in an ethnically homogenous population. More importantly, future research would also begin to identify potential mechanisms and confounders for this association in the general population. Some factors can be excluded, such as active smoking which was controlled for by excluding active smokers, but other factors will need to be tested for such as secondhand smoke, obesity, and/or exercise. As an example, obesity may be associated with temperature extremes as hot summers and cold winters may discourage exercise,[144] and obesity may also affect lung volumes.[145] As one limitation of our study was the use of cross-sectional lung function data, further research could examine longitudinal data. Similar to the above proposed research for CF, data that is linked to an individual's location over a lifetime could parse out the effects of specific geographic regions and or age-climate interactions, which may be more important in the general population, who may be more mobile than individuals with CF.

A third avenue of research could focus on other respiratory diseases. Although a larger effect size of temperature on baseline lung function was seen in individuals with cystic fibrosis versus the general population, it is unknown whether this would also hold true with the baseline lung function of other respiratory diseases. If a larger effect size was seen in some or all other respiratory diseases, this would imply that individuals with respiratory disease in general are more susceptible to the effects of temperature, and

hence climate change. Alternatively, if the effect size in other respiratory diseases is similar to that of the general population, that would imply that unique mechanisms are present in cystic fibrosis that alter outcomes in warmer temperatures. The presence of large multicenter disease-specific studies may lend itself to studying this question (e.g., the COPDGene study[146]).

Lastly, research could still continue to focus on gene-environment interactions in cystic fibrosis with ambient temperature. Although, we did not identify an interaction within the study population for the second part of the study, only a limited number of SNPs were tested, and a genome-wide study may identify an interaction(s). In addition, given that larger populations are frequently needed to identify gene-environment interactions compared to identifying simply loci of interest or an environmental exposure of interest,[128] more precise quantification of the environmental factor of temperature may be required to improve the power to detect an interaction. Specifically, instead of utilizing a mean annual ambient temperature as was done in this study, lung function could be linked to the temperature at the time lung function was performed, which could be done at least by month and year; this increased granularity may also provide better estimates of effect size for other temperature-related analyses.

An alternate approach to studying gene-environment interactions may be to utilize variable trajectories using mixed modelling approaches with multiple data points for both temperature and lung function. Although many traditional gene-environment interaction analyses have treated both the environmental factor(s) and outcome measure as a fixed measure obtained at a single time point, such an approach may not capture the full variability associated with a chronic disease with a declining health status or

environmental exposures that change over a lifetime. Equally importantly, genome-wide approaches that phenotype an individual with longitudinal data should increase statistical power over phenotyping an individual with data from a single time point.[147]

In conclusion, we previously found that warmer ambient temperatures are associated with lower lung function in cystic fibrosis, and have now observed evidence of partial mediation of this association through respiratory pathogens and also observed this association in the general population. Further research is required to confirm causality and determine mechanisms underlying this relationship as it may have increased ramifications with climate change.

REFERENCES

1. Riordan JR, Rommens JM, Kerem B, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 1989;245:1066-73.
2. Kerem E, Corey M, Kerem B-S, et al. The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (deltaF508). NEnglJMed 1990;323:1517-22.
3. Correlation between genotype and phenotype in patients with cystic fibrosis. The Cystic Fibrosis Genotype-Phenotype Consortium. NEnglJ Med 1993;329:1308-13.
4. Koch C, Cuppens H, Rainisio M, et al. European Epidemiologic Registry of Cystic Fibrosis (ERCF): comparison of major disease manifestations between patients with different classes of mutations. PediatrPulmonol 2001;31:1-12.
5. Gabriel SE, Brigman KN, Koller BH, Boucher RC, Stutts MJ. Cystic fibrosis heterozygote resistance to cholera toxin in the cystic fibrosis mouse model. Science 1994;266:107-9.
6. Modiano G, Ciminelli BM, Pignatti PF. Cystic fibrosis and lactase persistence: a possible correlation. EurJHumGenet 2007;15:255-9.
7. Poolman EM, Galvani AP. Evaluating candidate agents of selective pressure for cystic fibrosis. JR SocInterface 2007;4:91-8.
8. Courtney JM, Bradley J, Mccaughan J, et al. Predictors of mortality in adults with cystic fibrosis. PediatrPulmonol 2007;42:525-32.
9. Konstan MW, Morgan WJ, Butler SM, et al. Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. JPediatr 2007;151:134-9, 9.
10. Tablan OC, Martone WJ, Doershuk CF, et al. Colonization of the respiratory tract with Pseudomonas cepacia in cystic fibrosis. Risk factors and outcomes. Chest 1987;91:527-32.
11. Isles A, Maclusky I, Corey M, et al. Pseudomonas cepacia infection in cystic fibrosis: an emerging problem. JPediatr 1984;104:206-10.
12. Lipuma JJ. Update on the Burkholderia cepacia complex. CurrOpinPulmMed 2005;11:528-33.
13. Dasenbrook EC. Update on methicillin-resistant Staphylococcus aureus in cystic fibrosis. CurrOpinPulmMed 2011;17:437-41.
14. Dasenbrook EC, Checkley W, Merlo CA, Konstan MW, Lechtzin N, Boyle MP. Association between respiratory tract methicillin-resistant Staphylococcus aureus and survival in cystic fibrosis. JAMA 2010;303:2386-92.
15. Vanderhelst E, De ML, Verbanck S, Pierard D, Vincken W, Malfroot A. Prevalence and impact on FEV(1) decline of chronic methicillin-resistant Staphylococcus aureus (MRSA) colonization in patients with cystic fibrosis. A single-center, case control study of 165 patients. JCystFibros 2012;11:2-7.
16. Li Z, Kosorok MR, Farrell PM, et al. Longitudinal development of mucoid Pseudomonas aeruginosa infection and lung disease progression in children with cystic fibrosis. JAMA 2005;293:581-8.
17. Pritt B, O'Brien L, Winn W. Mucoid Pseudomonas in cystic fibrosis. AmJClinPathol 2007;128:32-4.
18. Rosenfeld M, Emerson J, Accurso F, et al. Diagnostic accuracy of oropharyngeal cultures in infants and young children with cystic fibrosis. Pediatr Pulmonol 1999;28:321-8.
19. Armstrong DS, Grimwood K, Carlin JB, Carzino R, Olinsky A, Phelan PD. Bronchoalveolar lavage or oropharyngeal cultures to identify lower respiratory pathogens in infants with cystic fibrosis. Pediatr Pulmonol 1996;21:267-75.
20. Sawicki GS, Sellers DE, Robinson WM. High treatment burden in adults with cystic fibrosis: challenges to disease self-management. Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society 2009;8:91-6.
21. Collaco JM, Green DM, Cutting GR, Naughton KM, Mogayzel PJ, Jr. Location and duration of treatment of cystic fibrosis respiratory exacerbations do not affect outcomes. AmJRespirCrit Care Med 2010;182:1137-43.
22. Accurso FJ, Rowe SM, Clancy JP, et al. Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. The New England journal of medicine 2010;363:1991-2003.
23. Davies JC, Wainwright CE, Canny GJ, et al. Efficacy and safety of ivacaftor in patients aged 6 to 11 years with cystic fibrosis with a G551D mutation. American journal of respiratory and critical care medicine 2013;187:1219-25.

24. De Boeck K, Munck A, Walker S, et al. Efficacy and safety of ivacaftor in patients with cystic fibrosis and a non-G551D gating mutation. Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society 2014;13:674-80.
25. Flume PA, Liou TG, Borowitz DS, et al. Ivacaftor in subjects with cystic fibrosis who are homozygous for the F508del-CFTR mutation. Chest 2012;142:718-24.
26. Wainwright CE, Elborn JS, Ramsey BW, et al. Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. The New England journal of medicine 2015.
27. Sanders DB, Bittner RC, Rosenfeld M, Redding GJ, Goss CH. Pulmonary exacerbations are associated with subsequent FEV1 decline in both adults and children with cystic fibrosis. PediatrPulmonol 2011;46:393-400.
28. Vanscoy LL, Blackman SM, Collaco JM, et al. Heritability of lung disease severity in cystic fibrosis. AmJRespirCrit Care Med 2007;175:1036-43.
29. Collaco JM, Blackman SM, McGready J, Naughton KM, Cutting GR. Quantification of the relative contribution of environmental and genetic factors to variation in cystic fibrosis lung function. JPediatr 2010;157:802-7.
30. Stanke F, Becker T, Kumar V, et al. Genes that determine immunology and inflammation modify the basic defect of impaired ion conductance in cystic fibrosis epithelia. JMedGenet 2010;48:24-31.
31. Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. NEnglJMed 1992;326:1187-91.
32. Schluchter MD, Konstan MW, Davis PB. Jointly modelling the relationship between survival and pulmonary function in cystic fibrosis patients. StatMed 2002;21:1271-87.
33. Collaco JM, Vanscoy L, Bremer L, et al. Interactions between secondhand smoke and genes that affect cystic fibrosis lung disease. JAMA 2008;299:417-24.
34. Beydon N, Amsallem F, Bellet M, et al. Pulmonary function tests in preschool children with cystic fibrosis. Am J RespirCrit Care Med 2002;166:1099-104.
35. 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.
36. Rubin BK. Exposure of children with cystic fibrosis to environmental tobacco smoke. NEnglJMed 1990;323:782-8.
37. Smyth A, O'Hea U, Williams G, Smyth R, Heaf D. Passive smoking and impaired lung function in cystic fibrosis. ArchDisChild 1994;71:353-4.
38. Goss CH, Newsom SA, Schildcrout JS, Sheppard L, Kaufman JD. Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis. AmJRespirCrit Care Med 2004;169:816-21.
39. Schechter MS, McColley SA, Silva S, Haselkorn T, Konstan MW, Wagener JS. Association of socioeconomic status with the use of chronic therapies and healthcare utilization in children with cystic fibrosis. JPediatr 2009;155:634-9.
40. 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 RespirCrit Care Med 2001;163:1331-7.
41. Schechter MS, Margolis PA. Relationship between socioeconomic status and disease severity in cystic fibrosis. J Pediatr 1998;132:260-4.
42. 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 RespirCrit Care Med 1997;155:1921-4.
43. Quittner AL, Schechter MS, Rasouliyan L, Haselkorn T, Pasta DJ, Wagener JS. Impact of socioeconomic status, race, and ethnicity on quality of life in patients with cystic fibrosis in the United States. Chest 2010;137:642-50.
44. Collaco JM, McGready J, Green DM, et al. Effect of Temperature on Cystic Fibrosis Lung Disease and Infections: A Replicated Cohort Study. PLoSOne 2011;6:e27784.
45. Patz JA, Campbell-Lendrum D, Holloway T, Foley JA. Impact of regional climate change on human health. Nature 2005;438:310-7.
46. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. Am J RespirCrit Care Med 1999;159:179-87.

47. Pittman JE, Calloway EH, Kiser M, et al. Age of Pseudomonas aeruginosa acquisition and subsequent severity of cystic fibrosis lung disease. PediatrPulmonol 2010.
48. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. JPersSocPsychol 1986;51:1173-82.
49. Kenny DA. Mediation. http://davidakennynet/cm/mediatehtm2009.
50. Kulich M, Rosenfeld M, Campbell J, et al. Disease-specific reference equations for lung function in patients with cystic fibrosis. AmJRespirCrit Care Med 2005;172:885-91.
51. Green DM, McDougal KE, Blackman SM, et al. Mutations that permit residual CFTR function delay acquisition of multiple respiratory pathogens in CF patients. RespirRes 2010;11:140.
52. Wright FA, Strug LJ, Doshi VK, et al. Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2. NatGenet 2011;43:539-46.
53. Chalmers JD, Fleming GB, Hill AT, Kilpatrick DC. Impact of mannose-binding lectin insufficiency on the course of cystic fibrosis: A review and meta-analysis. Glycobiology 2011;21:271-82.
54. Drumm ML, Konstan MW, Schluchter MD, et al. Genetic modifiers of lung disease in cystic fibrosis. NEnglJMed 2005;353:1443-53.
55. Dorfman R, Sandford A, Taylor C, et al. Complex two-gene modulation of lung disease severity in children with cystic fibrosis. JClinInvest 2008;118:1040-9.
56. Bremer LA, Blackman SM, Vanscoy LL, et al. Interaction between a novel TGFB1 haplotype and CFTR genotype is associated with improved lung function in cystic fibrosis. HumMolGenet 2008;17:2228-37.
57. Corvol H, Boelle PY, Brouard J, et al. Genetic variations in inflammatory mediators influence lung disease progression in cystic fibrosis. PediatrPulmonol 2008;43:1224-32.
58. Faria EJ, Faria IC, Ribeiro JD, Ribeiro AF, Hessel G, Bertuzzo CS. Association of MBL2, TGF-beta1 and CD14 gene polymorphisms with lung disease severity in cystic fibrosis. JBrasPneumol 2009;35:334-42.
59. Gu Y, Harley IT, Henderson LB, et al. Identification of IFRD1 as a modifier gene for cystic fibrosis lung disease. Nature 2009;458:1039-42.
60. Todd JL, Goldstein DB, Ge D, Christie J, Palmer SM. The state of genome-wide association studies in pulmonary disease: a new perspective. AmJRespirCrit Care Med 2011;184:873-80.
61. Wilk JB, Chen TH, Gottlieb DJ, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoSGenet 2009;5:e1000429.
62. Repapi E, Sayers I, Wain LV, et al. Genome-wide association study identifies five loci associated with lung function. NatGenet 2010;42:36-44.
63. Hancock DB, Eijgelsheim M, Wilk JB, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. NatGenet 2010;42:45-52.
64. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. AnnuRevGenomics HumGenet 2009;10:387-406.
65. Nair RP, Duffin KC, Helms C, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. NatGenet 2009;41:199-204.
66. Sosnay PR, Siklosi KR, Van Goor F, et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nature genetics 2013;45:1160-7.
67. Kerem E, Corey M, Kerem BS, et al. The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (delta F508). The New England journal of medicine 1990;323:1517-22.
68. Rowe SM, Heltshe SL, Gonska T, et al. Clinical mechanism of the cystic fibrosis transmembrane conductance regulator potentiator ivacaftor in G551D-mediated cystic fibrosis. American journal of respiratory and critical care medicine 2014;190:175-84.
69. O'Sullivan BP, Orenstein DM, Milla CE. Pricing for orphan drugs: will the market bear what society cannot? JAMA 2013;310:1343-4.
70. O'Connor GT, Quinton HB, Kneeland T, et al. Median household income and mortality rate in cystic fibrosis. Pediatrics 2003;111:e333-e9.
71. Stephenson A, Hux J, Tullis E, Austin PC, Corey M, Ray J. Socioeconomic status and risk of hospitalization among individuals with cystic fibrosis in Ontario, Canada. PediatrPulmonol 2011;46:376-84.

72. Ranganathan SC, Skoric B, Ramsay KA, et al. Geographical differences in first acquisition of Pseudomonas aeruginosa in cystic fibrosis. Annals of the American Thoracic Society 2013;10:108-14.
73. Psoter KJ, Rosenfeld M, De Roos AJ, Mayer JD, Wakefield J. Differential geographical risk of initial Pseudomonas aeruginosa acquisition in young US children with cystic fibrosis. American journal of epidemiology 2014;179:1503-13.
74. Kopp BTN, L.; Paul, G.; Tobias, J.; Ramanathan, C.; Hayes, D. Geographic variations in cystic fibrosis in the United States: An analysis of the U.S. Cystic Fibrosis Foundation Patient Registry. Pediatr Pulmonol 2014;49.
75. Hicks R, Tingley D. Causal mediation analysis. Stata J 2011;11:605-19.
76. Sobel ME. Asymptotic intervals for indirect effects in structural equations models. In: Leinhart S, ed. Sociological Methodology 1982. San Francisco, CA: Jossey-Bass; 1982:290-312.
77. Calculation for the Sobel Test. 2010-2015. (Accessed 12/24/2015, at http://quantpsy.org/sobel/sobel.htm.)
78. Lin MF, Lucas HC, Shmueli G. Too Big to Fail: Large Samples and the p-Value Problem. Inform Syst Res 2013;24:906-17.
79. Abdallah M, Benoliel C, Ferreira-Theret P, Drider D, Dhulster P, Chihib NE. Effect of culture conditions on the resistance of Pseudomonas aeruginosa biofilms to disinfecting agents. Biofouling 2015;31:49-59.
80. Abdallah M, Chataigne G, Ferreira-Theret P, et al. Effect of growth temperature, surface type and incubation time on the resistance of Staphylococcus aureus biofilms to disinfectants. Applied microbiology and biotechnology 2014;98:2597-607.
81. S C, SP G. Effects of growth temperature on polystyrene adhesion of Pseudomonas aeruginosa ATCC 27853. Brazilian Journal of Microbiology 2006;37:205-7.
82. De Soyza A, Hall AJ, Mahenthiralingam E, et al. Developing an international Pseudomonas aeruginosa reference panel. MicrobiologyOpen 2013;2:1010-23.
83. Grundmann H, Aanensen DM, van den Wijngaard CC, et al. Geographic distribution of Staphylococcus aureus causing invasive infections in Europe: a molecular-epidemiological analysis. PLoS medicine 2010;7:e1000215.
84. Adjemian J, Olivier KN, Prevots DR. Nontuberculous mycobacteria among patients with cystic fibrosis in the United States: screening practices and environmental risk. American journal of respiratory and critical care medicine 2014;190:581-6.
85. Bezirtzoglou C, Dekas K, Charvalos E. Climate changes, environment and infection: facts, scenarios and growing awareness from the public health community within Europe. Anaerobe 2011;17:337-40.
86. Collaco JM, Cutting GR. Update on gene modifiers in cystic fibrosis. Current opinion in pulmonary medicine 2008;14:559-66.
87. McDougal KE, Green DM, Vanscoy LL, et al. Use of a modeling framework to evaluate the effect of a modifier gene (MBL2) on variation in cystic fibrosis. EurJHumGenet 2010;18:680-4.
88. Gauderman W, Morrison J. QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies, http://hydra.usc.edu/gxe. 2006.
89. Hukkinen M, Kaprio J, Broms U, et al. Heritability of lung function: a twin study among never-smoking elderly women. Twin research and human genetics : the official journal of the International Society for Twin Studies 2011;14:401-7.
90. Ghio AJ, Crapo RO, Elliott CG, et al. Heritability estimates of pulmonary function. Chest 1989;96:743-6.
91. McClearn GE, Svartengren M, Pedersen NL, Heller DA, Plomin R. Genetic and environmental influences on pulmonary function in aging Swedish twins. Journal of gerontology 1994;49:264-8.
92. Zhai G, Valdes AM, Cherkas L, Clement G, Strachan D, Spector TD. The interaction of genes and smoking on forced expiratory volume: a classic twin study. Chest 2007;132:1772-7.
93. Organization WH. WHO Guidelines for Indoor Air Quality. Household Fuel Combustion. Geneva, Switzerland: WHO Document Production Services; 2014.
94. Moritsugu KP. The 2006 Report of the Surgeon General: the health consequences of involuntary exposure to tobacco smoke. American journal of preventive medicine 2007;32:542-3.

95. Czogala J, Goniewicz ML, Fidelus B, Zielinska-Danch W, Travers MJ, Sobczak A. Secondhand exposure to vapors from electronic cigarettes. Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco 2014;16:655-62.
96. Rice MB, Rifas-Shiman SLM, Litonjua AA, et al. Lifetime Exposure to Ambient Pollution and Lung Function in Children. American journal of respiratory and critical care medicine 2015.
97. Rice MB, Ljungman PL, Wilker EH, et al. Long-term exposure to traffic emissions and fine particulate matter and lung function decline in the Framingham heart study. American journal of respiratory and critical care medicine 2015;191:656-64.
98. Adam M, Schikowski T, Carsin AE, et al. Adult lung function and long-term air pollution exposure. ESCAPE: a multicentre cohort study and meta-analysis. The European respiratory journal 2015;45:38-50.
99. Bernstein AS, Rice MB. Lungs in a warming world: climate change and respiratory health. Chest 2013;143:1455-9.
100. Hansel NN, McCormack MC, Kim V. The Effects of Air Pollution and Temperature on COPD. Copd 2015:1-8.
101. Li S, Baker PJ, Jalaludin BB, et al. An Australian national panel study of diurnal temperature range and children's respiratory health. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology 2014;112:348-53 e1-8.
102. Mireku N, Wang Y, Ager J, Reddy RC, Baptist AP. Changes in weather and the effects on pediatric asthma exacerbations. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology 2009;103:220-4.
103. Wang X, Dockery DW, Wypij D, Fay ME, Ferris BG, Jr. Pulmonary function between 6 and 18 years of age. PediatrPulmonol 1993;15:75-88.
104. Specifying Weighting Parameters (National Center for Health Statistics). (Accessed 12/21/2015, at http://www.cdc.gov/nchs/tutorials/nhanes/SurveyDesign/Weighting/intro_iii.htm.)
105. Harris RJ, Bradburn MJ, Deeks JJ, Harbord RM, Altman DG, Sterne JAC. metan: fixed- and random-effects meta-analysis. Stata J 2008;8:3-28.
106. Shi L, Kloog I, Zanobetti A, Liu P, Schwartz JD. Impacts of Temperature and its Variability on Mortality in New England. Nature climate change 2015;5:988-91.
107. D'Amato G, Vitale C, De Martino A, et al. Effects on asthma and respiratory allergy of Climate change and air pollution. Multidisciplinary respiratory medicine 2015;10:39.
108. Voss JD, Masuoka P, Webber BJ, Scher AI, Atkinson RL. Association of elevation, urbanization and ambient temperature with obesity prevalence in the United States. International journal of obesity 2013;37:1407-12.
109. Rudan I, O'Brien KL, Nair H, et al. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. Journal of global health 2013;3:010401.
110. Ehteshami-Afshar S, FitzGerald JM, Doyle-Waters MM, Sadatsafavi M. The global economic burden of asthma and chronic obstructive pulmonary disease. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease 2016;20:11-23.
111. Collaborators GBDRF, Forouzanfar MH, Alexander L, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:2287-323.
112. Akinbami LJ, Simon AE, Rossen LM. Changing Trends in Asthma Prevalence Among Children. Pediatrics 2016;137:1-7.
113. Lavinghouze SR, Malarcher A, Jama A, Neff L, Debrot K, Whalen L. Trends in Quit Attempts Among Adult Cigarette Smokers - United States, 2001-2013. MMWR Morbidity and mortality weekly report 2015;64:1129-35.
114. Stocker TFQ, D. Climate Change 2013: The Physical Science Basis. New York, NY2013.
115. Costello A, Abbas M, Allen A, et al. Managing the health effects of climate change: Lancet and University College London Institute for Global Health Commission. Lancet 2009;373:1693-733.
116. Rosenfeld M, VanDevanter DR, Ren CL, et al. Decline in lung function does not predict future decline in lung function in cystic fibrosis patients. Pediatr Pulmonol 2015;50:856-62.

117. Barr HL, Britton J, Smyth AR, Fogarty AW. Association between socioeconomic status, sex, and age at death from cystic fibrosis in England and Wales (1959 to 2008): cross sectional study. BMJ 2011;343:d4662.
118. Eakin MN, Bilderback A, Boyle MP, Mogayzel PJ, Riekert KA. Longitudinal association between medication adherence and lung health in people with cystic fibrosis. JCystFibros 2011;10:258-64.
119. Hebestreit H, Kieser S, Junge S, et al. Long-term effects of a partially supervised conditioning programme in cystic fibrosis. EurRespirJ 2010;35:578-83.
120. Kriemler S, Kieser S, Junge S, et al. Effect of supervised training on FEV in cystic fibrosis: A randomised controlled trial. JCystFibros 2013.
121. Moorcroft AJ, Dodd ME, Morris J, Webb AK. Individualised unsupervised exercise training in adults with cystic fibrosis: a 1 year randomised controlled trial. Thorax 2004;59:1074-80.
122. Schneiderman-Walker J, Pollock SL, Corey M, et al. A randomized controlled trial of a 3-year home exercise program in cystic fibrosis. JPediatr 2000;136:304-10.
123. Selvadurai HC, Blimkie CJ, Meyers N, Mellis CM, Cooper PJ, Van Asperen PP. Randomized controlled study of in-hospital exercise training programs in children with cystic fibrosis. PediatrPulmonol 2002;33:194-200.
124. Schneiderman JE, Wilkes DL, Atenafu EG, et al. Longitudinal relationship between physical activity and lung health in patients with cystic fibrosis. The European respiratory journal 2014;43:817-23.
125. Devore S, Champion RW. Driving population health through accountable care organizations. Health affairs 2011;30:41-50.
126. Mogayzel PJ, Jr., Naureckas ET, Robinson KA, et al. Cystic Fibrosis Foundation pulmonary guideline. pharmacologic approaches to prevention and eradication of initial Pseudomonas aeruginosa infection. Annals of the American Thoracic Society 2014;11:1640-50.
127. Aschard H, Lutz S, Maus B, et al. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Human genetics 2012;131:1591-613.
128. Murcray CE, Lewinger JP, Conti DV, Thomas DC, Gauderman WJ. Sample size requirements to detect gene-environment interactions in genome-wide association studies. GenetEpidemiol 2011;35:201-10.
129. Eschenbacher WL, Moore TB, Lorenzen TJ, Weg JG, Gross KB. Pulmonary responses of asthmatic and normal subjects to different temperature and humidity conditions in an environmental chamber. Lung 1992;170:51-62.
130. Martin SL, Kirkner GJ, Mayo K, Matthews CE, Durstine JL, Hebert JR. Urban, rural, and regional variations in physical activity. The Journal of rural health : official journal of the American Rural Health Association and the National Rural Health Care Association 2005;21:239-44.
131. Singh GK, Kogan MD, Siahpush M, van Dyck PC. Prevalence and correlates of state and regional disparities in vigorous physical activity levels among US children and adolescents. Journal of physical activity & health 2009;6:73-87.
132. von Glasow R, Jickells TD, Baklanov A, et al. Megacities and large urban agglomerations in the coastal zone: interactions between atmosphere, land, and marine ecosystems. Ambio 2013;42:13-28.
133. Community and Hospital-Associated MRSA. (Accessed 1/29/2016, 2015, at http://www.cddep.org/projects/resistance_map/community_and_hospital_associated_mrsa.)
134. Proportion of Methicillin Resistant Staphylococcus aureus (MRSA) Isolates in Participating Countries in 2014. (Accessed 1/29/2016, 2016, at http://ecdc.europa.eu/en/healthtopics/antimicrobial_resistance/database/Pages/map_reports.aspx.)
135. Centers for Disease C, Prevention. Chronic obstructive pulmonary disease among adults--United States, 2011. MMWR Morbidity and mortality weekly report 2012;61:938-43.
136. Li S, Baker PJ, Jalaludin BB, Marks GB, Denison LS, Williams GM. Ambient temperature and lung function in children with asthma in Australia. The European respiratory journal 2014;43:1059-66.
137. Discher DP, Palmer A, Hibdon G, Drizd TA. Computer-assisted spirometry data analysis for the National Health and Nutrition Examination Survey, 1971-80. Vital and health statistics Series 2, Data evaluation and methods research 1980:1-48.

138. Liu X, O'Donnell N, Landstrom A, Skach WR, Dawson DC. Thermal instability of DeltaF508 cystic fibrosis transmembrane conductance regulator (CFTR) channel function: protection by single suppressor mutations and inhibiting channel activity. Biochemistry 2012;51:5113-24.
139. Wang W, Okeyo GO, Tao B, Hong JS, Kirk KL. Thermally unstable gating of the most common cystic fibrosis mutant channel (DeltaF508): "rescue" by suppressor mutations in nucleotide binding domain 1 and by constitutive mutations in the cytosolic loops. The Journal of biological chemistry 2011;286:41937-48.
140. McKone EF, Emerson SS, Edwards KL, Aitken ML. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. Lancet 2003;361:1671-6.
141. Masica DL, Sosnay PR, Raraigh KS, Cutting GR, Karchin R. Missense variants in CFTR nucleotide-binding domains predict quantitative phenotypes associated with cystic fibrosis disease severity. Human molecular genetics 2015;24:1908-17.
142. Rose DP, Ratterman ME, Griffin DK, et al. Adhesive RFID Sensor Patch for Monitoring of Sweat Electrolytes. IEEE transactions on bio-medical engineering 2015;62:1457-65.
143. Gao W, Emaminejad S, Nyein HY, et al. Fully integrated wearable sensor arrays for multiplexed in situ perspiration analysis. Nature 2016;529:509-14.
144. von Hippel P, Benson R. Obesity and the natural environment across US counties. American journal of public health 2014;104:1287-93.
145. Jones RL, Nzekwu MM. The effects of body mass index on lung volumes. Chest 2006;130:827-33.
146. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. Copd 2010;7:32-43.
147. Furlotte NA, Eskin E, Eyheramendy S. Genome-wide association mapping with longitudinal data. Genetic epidemiology 2012;36:463-71.

CURRICULUM VITAE

J. Michael Collaco was born in Oregon, Ohio in 1973. He completed a B.S. in Biochemistry/Anthropology at Yale University (New Haven, CT), an M.D. and M.S. in Anatomy at Case Western Reserve University (Cleveland, OH), and Pediatrics Residency, Pediatric Pulmonology Fellowship, an M.B.A., and an M.P.H. at Johns Hopkins Unversity (Baltimore, MD). He has been on faculty at the Johns Hopkins University School of Medicine since 2008 with research interests in the epidemiology of pediatric chronic respiratory diseases.