Imperial College London

Imperial College London

Faculty of Medicine

School of Public Health

Department of Epidemiology and Biostatistics

Role of environmental factors on cystic fibrosis disease severity: Spatio-temporal analyses

By

Muhammad Saleem Khan

A thesis submitted in part fulfilment of the requirements for the degree of Doctor of

Philosophy of Imperial College London

July 2023

Declaration of originality

I declare that this thesis titled ***Role of environmental factors on cystic fibrosis disease severity: Spatio-temporal analyses*** *i*s entirely my own work and that any material that could be construed as the work of others is properly cited and referenced, and/or due acknowledgement is given.

Name of student: Muhammad Saleem Khan

Name of supervisors: Dr Frédéric B. Piel

Prof Nicholas Simmonds

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Cystic fibrosis (CF) is a severe respiratory illness, and environmental factors are increasingly recognized as potential contributors to its severity. This PhD research investigated the impact of air pollution and bioaerosols on the health of people with CF (pwCF) in the UK. Limited research existed on this topic, despite evidence linking environmental exposures to respiratory health.

Using data from the UK CF Registry – representing over 99% of pwCF in the UK – the research employed spatio-temporal analyses to explore environmental exposures and disease severity. The primary areas of focus were:

- Short- and long-term effects of air pollutants $PM_{2.5}$ and NO_2 on lung function and pulmonary exacerbations. Results suggested a significant association between $PM_{2.5}$ exposure and increased intravenous antibiotic use in adult CF patients, while the impact on lung function remained inconclusive.
- Potential health risks associated with living near industrial composting sites. This novel study found a possible link between proximity to composting facilities and reduced lung function, as well as increased exacerbation rates in pwCF.

These findings highlight the potential dangers posed by air pollution and bioaerosols to the health of pwCF. The research underscores the importance of:

- Public health policies aimed at reducing air pollution, particularly $PM_{2.5}$ levels.
- Strategic siting of composting facilities to minimize potential health risks for vulnerable populations.
- Collaboration between healthcare providers and pwCF to understand residential environments and implement mitigation strategies.

Further research is needed to explore the mechanisms behind these associations and confirm the observed effects. Ultimately, this work aims to contribute to improved public health policies and environmental management, leading to better health outcomes for pwCF

Summary of the project

There is growing evidence worldwide about the impact of environmental and socio-economic status on human health (Lago-Peñas *et al.*, 2021; Vidal *et al.*, 2022). Subjects with respiratory diseases such as chronic obstructive pulmonary disease (COPD) and asthma are particularly vulnerable to the negative effects of air pollution (Jiang *et al.*, 2016; Tiotiu *et al.*, 2020; Wang *et al.*, 2022a). Nevertheless, there is a clear lack of evidence in the UK on the role of air quality on the severity of cystic fibrosis (CF) (Blayac *et al.*, 2022; Brugha *et al.*, 2018). Many pollutants contribute to air quality. Particulate matter ($PM_{2.5}$ and PM_{10}) and NO_2 tend to be the most commonly studied due to their major impact on health (Schraufnagel *et al.*, 2019a; Schraufnagel *et al.*, 2019b). Although the role of other pollutants, such as bioaerosols, is less clear, they are likely to play an important role in bacterial and fungal infections in people with CF (pwCF) and other respiratory diseases (Roca-Barcelo *et al.*, 2020; Samake *et al.*, 2017).

The primary aim of my PhD research was to address the important research gap in our knowledge of the impact of air pollutants and bioaerosols on disease severity in pwCF in the UK, building on the limited evidence from the US and some European countries. I believe that by contributing to uncover the complex interactions underlying disease severity, my research would offer valuable insights for public health strategies and environmental management, that could have direct benefits for pwCF.

Throughout my thesis, I used data from the UK CF Registry (2007-2017), which includes data from every specialist centre and clinic across the UK and represents over 99% of pwCF. I focused on spatio-temporal analyses based on small-area quantitative methods to study environmental-health associations This project, conducted in the UK Small-Area Health Statistics Unit (SAHSU) (SAHSU, 2024), was funded by the UK CF Trust (https://www.cysticfibrosis.org.uk/) and The Medical Research Council (MRC) Doctoral Training Partnership at Imperial (Imperial College London, 2024), in collaboration with the CF-EpiNet Strategic Research Centre (Cystic Fibrosis Trust, 2024).

To begin, I wanted to assess the short-term (one year) and long-term (ten years) influence of $PM_{2.5}$ and NO_2 on disease severity in CF. For this, I used two key measures commonly used by clinicians to assess disease severity in CF: i) percent predicted forced expiratory volume in one second ($ppFEV_1$) as a measure of lung function, and ii) the annual number of days on

intravenous antibiotics for pulmonary exacerbation (#IVdays). To explore the interactions between bioaerosol exposures and health in pwCF, I decided to focus on industrial composting sites. Composting is a growing element of sustainable waste management approaches across the UK, and commercial-scale composting facilities can generate large amounts of bioaerosols, which are relevant to fungal and bacterial infections and human health (Robertson *et al.*, 2019). A more detailed summary of each one of these studies is provided below:

1. Short-term effect of $PM_{2.5}$ and NO_2 on disease severity in pwCF in Great Britain (GB).

Rationale: The study aims to address a knowledge gap in the short-term (one-year) impact of $PM_{2.5}$ and NO_2 on disease severity in pwCF in GB. Understanding how these pollutants contribute to disease progression and severity in the CF population could help develop more efficient ways and technologies for reducing complications associated with poor air quality.

Methods: I used a semi-individual cross-sectional study design combining high-quality health data from the 2016 UK CF Registry and residential postcode exposure estimates from high-resolution (25x25m) land-use regression models of $PM_{2.5}$ and NO_2 and (2015) concentrations in GB. Linear and zero-inflated negative binomial regression adjusted for age, sex, weight, height, genotype, *Pseudomonas aeruginosa* (*PA*) and deprivation (Townsend index) were used to explore the association between exposures ($PM_{2.5}$ and NO_2) quintiles and outcomes ($ppFEV_1$ & #IVdays).

Results: A total of 8,965 pwCF were included: mean age: 20.1 (SD =14.0) years; 53.3% male; 49.2% homozygous *F508del;* 44.8% positive for *PA*. Mean (SD) $ppFEV_1$ and #IVdays were 70.6 (24) and 16.7 (28.5) days, respectively. Annual mean concentrations ranged between 1.6 and 71.2 $\mu g/m^3$ for NO_2 (WHO guideline: 40 $\mu g/m^3$), and 3.2 and 14.0 $\mu g/m^3$ for $PM_{2.5}$ (WHO guideline: 10 $\mu g/m^3$). In CF adults, I did not find a clear relationship between $ppFEV_1$ and either NO_2 (0.38%: 95%CI; -1.71, 2.46, p=0.724) or $PM_{2.5}$ (0.35%: 95%CI; -1.72, 2.42, p=0.740). There was a statistically significant association between $PM_{2.5}$ and #IVdays in adults with CF. The #IVdays per year among all adults with CF in the Q5 quintile of $PM_{2.5}$ increased by 1.11 (95% CI: 1.01, 1.21, p=0.036) and 1.08 (95% CI: 0.89, 1.17, p= 0.087), respectively before and after adjustments. The odds of experiencing zero IVdays were lower in the higher quintiles of $PM_{2.5}$ in the zero-inflated model, albeit this difference was not statistically significant in the fifth quintile (Q5) of $PM_{2.5}$ (1.07, 95%CI: 0.87, 1.32, p= 0.506). In children with CF, there was no

evidence of an association between $PM_{2.5}$ and #IVdays (0.98, 95%CI: 0.85, 1.11, p= 0.742). Although $ppFEV_1$ was lower in the higher quintile of NO_2 (-0.70, 95%CI: -3.05; 1.64, p= 0.556). and $PM_{2.5}$ (-1.73, 95%CI; -4.03, 0.58, p= 0.141). when compared to the lowest quintile (Q1) this difference did not reach statistical significance in children.

Conclusions: The study highlighted a significant association between $PM_{2.5}$ exposure and increased #IVdays in adults with CF, even after adjusting for covariates. However, this association was not present in the paediatric CF population. Moreover, the impact of air pollutants, specifically NO_2 and $PM_{2.5}$, on pulmonary function ($ppFEV_1$) in both adults and children with CF remained uncertain based on my findings.

Public health implications: This study suggests that reducing air pollution, particularly $PM_{2.5}$ could be an important strategy for preventing exacerbations and improving health outcomes in pwCF. Further research is needed to explore the reasons for the observed differences between adults and children and to confirm the impact of air pollution on lung function in CF. Public health policies aimed at reducing air pollution exposure would not only benefit pwCF but also the general population.

2. Long-term association between $PM_{2.5}$ exposure and lung function in pwCF in Greater London: a longitudinal study using the UK Cystic Fibrosis Registry

Rationale: While the impact of $PM_{2.5}$ on health in CF remains uncertain, there is a notable gap in the literature on its long-term effects on lung function decline in CF populations in the UK. This longitudinal study aimed to address this research gap by investigating the association between $PM_{2.5}$ and $ppFEV_1$ decline in pwCF living in London. The findings of this study could potentially guide clinical practice, public health initiatives, and ultimately lead to improved respiratory health and overall well-being for pwCF in London.

Methods: A longitudinal cohort study was used to evaluate the association between $PM_{2.5}$ exposure and the rate of decline in $ppFEV_1$ among pwCF in the Greater London Area (GLA). The study included pwCF with a residential postcode in the GLA in 2017 and who were at least six years old in 2008. A high-resolution (20m x 20m) model of $PM_{2.5}$ from the "London Air Pollution Toolkit" was used to assign residential exposure at the postcode level. Linear mixed models with random intercepts were used to estimate the excess decline in $ppFEV_1$ in pwCF

with higher long-term exposure compared to those with low exposure over the ten years of the study (2008-2017).

Results: A total of 393 pwCF (271 adults and 122 children) with 3,333 $ppFEV_1$ measurements were included. Over 43% of pwCF resided in the most deprived areas of the GLA. The mean $ppFEV_1$ decreased by 10% from 2008 to 2017, corresponding to an approximate 1% annual decline. In 2017, pwCF with higher $PM_{2.5}$ exposure (≥ 14.11 $\mu g/m^3$) experienced a mean $ppFEV_1$ decline of 14.5% (95% CI: 7.04%; 22.0%), while those with lower exposure (≤ 13.39 $\mu g/m^3$) presented a decline of 7.56% (95% CI: 4.10%; 11.00%).

Conclusions: High long-term exposure to $PM_{2.5}$ was associated with a statistically significant higher decline in $ppFEV_1$ among pwCF residing in London between 2008 and 2017. These findings highlight the potential detrimental impact of $PM_{2.5}$ on respiratory function in pwCF. Similar analyses need to be conducted for other pollutants including NO_2.

Public health implications: The study underscored the urgent need for further interventions addressing $PM_{2.5}$ exposure among pwCF in London. Minimising air pollution, especially $PM_{2.5}$, could potentially slow the decline in lung function, enhancing the respiratory health of this vulnerable CF population. Public health initiatives focusing on reducing $PM_{2.5}$ levels may contribute to improved clinical outcomes, emphasizing the importance of implementing air quality measures to safeguard the well-being of pwCF in urban settings like the GLA.

3. Investigating the influence of proximity to Registered Composting Sites (RCS) on respiratory health in pwCF in the UK.

Rationale: Understanding environmental risk factors that contribute to disease severity is crucial for improving health outcomes in pwCF. Composting sites emit various organic dust and bioaerosols, which have been linked to respiratory issues in the general population. However, limited research explored the potential impact of proximity to such sites on the respiratory health of pwCF. This study addresses this gap by investigating the association between residential proximity to RCS and lung function and pulmonary exacerbation (PEx) in pwCF across the UK.

Methods: I adopted a cross-sectional ecological small-area study design using 2016 data from the UK CF Registry. All RCS were geocoded using site postcodes provided by Environmental Agencies operating within the UK, namely the UK Environment Agency (England), Natural

Resources Wales, the Scottish Environment Protection Agency, and the Northern Ireland Environment Agency. PwCF's exposure was assigned based on residential postcodes. Three proxies of disease severity were used: $ppFEV_1$, #IVdays and fungal infection. Linear and Poisson regressions adjusted for age, gender, BMI, genotype, urban-rural classification and deprivation (Townsend index quintiles) were used to explore the association between health outcomes ($ppFEV_1$, #IVdays and fungal infection) and the distance within and beyond 4km around RCS.

Results: There were 256 active RCSs in the UK (209 in England; 24 in Wales; 21 in Scotland; and 2 in Northern Ireland) and 9,361 pwCF (children and adults) had an annual review in 2016. The average age of pwCF was 20.1 (*SD*=14.1) years, 53.3% were male and 49.2% were homozygous F508del. 19.2% lived in rural areas. Over 10% of pwCF (n= 1,015) lived within 4km of a RCS and had a mean $ppFEV_1$ of 68.0%, compared to 70.9% among those who lived beyond 4km. The mean (*SD*) #IVdays was 18.5 (30.0) per year. I found a 1.07% (95% CI, -2.29, 0.16) decrease in $ppFEV_1$ and 1.02 (95% CI, 1.01%, 1.04%) increase in #IVdays per year for PwCF residing within 4 km of a RCS compared to those living further away. There was no significant difference in fungal infection rates within and beyond 4 km of a RCS.

Conclusions: This novel study of pwCF using UK CF Registry data suggested that living close to a RCS may be associated with a decrease in $ppFEV_1$ and an increase in #IVdays. Further studies accounting for aerosol dispersion and fungal infections are recommended. If confirmed, these results may have important implications for the living environment of pwCF.

Public health implications: Overall, this study underscored the importance of considering environmental factors in the management of CF and highlighted the need for further research to safeguard the health of pwCF. If these initial findings are confirmed, the implications for public health policy and CF management are significant. Policymakers should prioritise careful siting of composting facilities, considering potential health risks, especially in areas with high CF populations. Implementing buffer zones and raising community awareness about potential exposure risks and mitigation strategies are crucial. Healthcare providers can work with pwCF to understand their residential environment and potential bioaerosol risks. These combined efforts can optimise living environments for people with CF and safeguard their health.

In summary, my doctoral thesis adds to the small body of evidence on environmental factors influencing CF disease severity. My research highlights the potential impact of both air pollutants and bioaerosols and the need for a clearer understanding of exposure sources. While my findings underscored the importance of preventive measures like air quality regulations and individual exposure reduction strategies, the complexity of environmental exposures must be considered. Future research should focus on understanding the effects of multiple pollutants, both indoors and outdoors, on CF health. By building on these findings, comprehensive public health measures that target not only outdoor air pollution but also potential indoor air quality concerns for pwCF can be developed. Such an approach holds promise for significantly improving the well-being of those with CF and other respiratory illnesses (Rawat & Kumar, 2023).

Acknowledgements

I would like to sincerely thank my supervisors, Dr Frédéric B. Piel and Prof Nicholas Simmonds, for their invaluable support and guidance throughout the journey of completing my PhD thesis. My study has been greatly shaped and improved by their knowledge, persistence, and dedication. I sincerely appreciate all of their support, wise criticism, and unshakable faith in my talents. I owe Dr. Frédéric B. Piel a debt of gratitude for his constant support, guidance, and inspiration. Sincere thanks for being such a great supervisor and for making my thesis possible.

I would like to express my sincere thanks to the Medical Research Council (MRC) Doctoral Training Partnership (DTP) and the CF Trust for their generous funding for my PhD studies. The financial assistance provided by the MRC DTP and CF Trust was crucial to the successful completion of my thesis research work. I am also grateful to both institutions for providing me with an additional three months of funding during my interruption of studies. This support helped me overcome the challenges I faced and allowed me to continue my research with renewed focus and determination.

I sincerely appreciate all of the support and encouragement the MRC DTP, CF Trust, and the School of Public Health have given me while I've pursued my academic goals. They have made significant contributions that have helped my thesis be completed successfully, and I am grateful for their constant support.

I am grateful to the UK Cystic Fibrosis Trust and the Registry Steering Committee for access to anonymised UK Cystic Fibrosis Registry data, and to the UK Environment Agency, Natural Resources Wales, the Scottish Environment Protection Agency and the Northern Ireland Environment Agency for sharing data on composting sites.

I want to express my sincere gratitude to Dr Benjamin Barratt for providing me with the air pollutants exposure data for Greater London and to Dr Weiyi Wang for sharing the NO_2 and $PM_{2.5}$ exposure model surfaces for Great Britain. These contributions, which have been essential to advancing my knowledge of air pollution and its effect on disease severity in pwCF, have significantly improved my PhD thesis.

I also thank Hima Daby, Gajanan Natu, Eric Johnson and Rob Elliott of the SAHSU database team for technical support.

I thank Iqbal Azam, Joel Heller, Brandon L. Parkes and Georges Bucyibaruta for their valuable statistical and technical support. Their knowledge and guidance have been essential to the accomplishment of my PhD work.

Finally, I want to express my gratitude to my family and friends for their love and support. Without their assistance, I could not have accomplished this PhD.

For my wife, kids

and parents!!!

Table of Contents

Summary of the project 5

Acknowledgements 11

Table of Contents 14

List of Tables 18

List of Figures 21

Abbreviations 25

Structure of the Project 27

1 Chapter I: General Introduction 29

1.1 Cystic Fibrosis 29

1.2 Disease classification 29

1.3 Clinical manifestations and disease progression 30

1.4 Infections in pwCF 31

1.5 Environmental factors 32

1.5.1 Air pollution 33

1.5.2 Impact of air pollution on health 34

1.5.3 Nitrogen dioxide (NO_2) 35

1.5.4 Particulate matter (PM) 38

1.5.5 Existing evidence on the relationship between outdoor air pollution and outcomes in pwCF 41

1.5.6 Bioaerosols 45

1.6 Hypothesis 46

1.7 Aims and objectives 46

2 Chapter II: Long-term (ten years) exposure to air pollutants and disease severity in the Greater London Area 48

2.1 Background 48

2.1.1 Impact of air pollution (exposures) on CF 48

2.1.2 Health outcomes in CF 50

2.1.3 Air pollution and CF lung function ($ppFEV_1$) 50

2.1.4 Air pollution and CF pulmonary exacerbations (PEx) 50

2.1.5 The rationale of the study 52

2.1.6 Hypothesis 52

2.1.7 Objective of the study 52

2.2 Materials and methods 52

2.2.1 Study design 52

2.2.2 Study geography 53
2.2.3 Study population 54
2.2.4 Exclusion criteria 54
2.2.5 Outcome variable 54
2.2.6 Lung function 54
2.2.7 Pulmonary exacerbation 55
2.2.8 Exposure variables and covariates 55
2.2.9 Air quality 56
2.2.10 Covariates 59
2.2.11 Spatial and statistical analysis 68
2.2.12 Ethical considerations 71
2.3 Results 71
2.3.1 Overview of the CF data 71
2.3.2 Correlation between outcome variables and air pollutants 77
2.3.3 Association between ppFEV$_1$ and covariates in children and adults 80
2.3.4 Lung function and air pollutants (NO_2 and $PM_{2.5}$) 82
2.3.5 #IVdays and air pollutants (NO_2 and $PM_{2.5}$) 86
2.3.6 Sensitivity analyses 92
2.4 Discussion 94
2.4.1 Lung function (ppFEV$_1$) 94
2.4.2 Pulmonary exacerbations (PEx) 99
2.4.3 Strengths and limitations of my study 105
2.4.4 Conclusion 108
2.4.5 Recommendations 108
3 Chapter III: Long-term (ten years) exposure to air pollutants and disease severity in the Greater London Area 110
3.1 Background 110
3.1.1 Research question 111
3.1.2 Hypothesis 111
3.1.3 Objective of the study 111
3.2 Materials and methods 111
3.2.1 Study design 111
3.2.2 Study geography and population 111
3.2.3 Dummy postcode 113
3.2.4 Outcome variable: Lung function (ppFEV$_1$) 114
3.2.5 Exposure variable 114

3.2.6 Covariates 115
3.2.7 Spatial and Statistical analysis 115
3.3 Results 118
3.3.1 Characteristics of pwCF for this analysis 120
3.3.2 Summary statistics of ppFEV$_1$ and PM$_{2.5}$ 123
3.3.3 Correlation within ppFEV$_1$ and PM$_{2.5}$ 126
3.3.4 Longitudinal analysis of PM$_{2.5}$ and ppFEV$_1$ 126
3.3.5 Sensitivity /Stratified analysis 131
3.4 Discussion 144
3.4.1 Lung function (ppFEV$_1$) 145
3.4.2 Lung function (ppFEV$_1$) and PM$_{2.5}$ 145
3.4.3 Sensitivity analyses 149
3.4.4 Strengths of the study 151
3.4.5 Limitations of the study 152
3.4.6 Conclusions 153
4 Chapter IV: Bioaerosol exposure and disease severity in the UK 154
4.1 Background 155
4.1.1 Hypothesis 158
4.1.2 The objective of the study 158
4.2 Materials and methods 159
4.2.1 Study design, population and the outcome variable 159
4.2.2 Exposure variable 160
4.2.3 Covariates 161
4.2.4 Statistical analysis 161
4.2.5 Ethics and approval 161
4.3 Results 162
4.3.1 Lung function 168
4.3.2 PEx 171
4.3.3 Fungal and bacterial infections 171
4.4 Discussion 173
4.4.1 Fungal infection 175
4.4.2 Strengths and limitations of the study 176
4.4.3 Conclusions and recommendations 177
5 Chapter V: Overall discussion 179
5.1.2 Strength of the evidence on the relationship between air pollution exposure, bioaerosol and CF outcomes 184

5.1.3 Recommendations for the CF community 185
5.1.4 Key public health implications of my findings 186
5.1.5 Implications for clinicians and pwCF 187
5.1.6 Directions for the future work 189
5.1.7 Conclusions 192
6 References 194
7 Appendices 219
7.1 Chapter II: Boxplots of the relationship between disease severity (ppFEV$_1$ and PEx) and air pollutants (NO_2 and $PM_{2.5)}$ in pwCF in GB 219
7.2 Chapter III Correlation plot 241

List of Tables

Table 2.1: Classification of *CFTR* genotypes with frequencies among pwCF (n=9,361) in the UK in 2016 60

Table 2.2: List of minimal function mutations 62

Table 2.3: List of RF mutations 64

Table 2.4: Summary statistics of the outcome variable and air pollutants for pwCF with a residential postcode in Great Britain in 2016 on the UK CF Registry 73

Table 2.5: Demographic characteristics of pwCF with residential postcodes of GB 75

Table 2.6: Correlation between air pollutants and key outcome variables among pwCF with a residential postcode of GB 77

Table 2.7: Univariate linear regressions between $ppFEV_1$ and characteristics of children and adults with CF in the UK CF Registry in GB 81

Table 2.8: Distribution of the summary statistics of #IV day for PEx in pwCF 92

Table 2.9: Linear regression between air pollutants (NO_2 and $PM_{2.5}$) and lung function in adults (age ≥ 16) by mutation (severe vs mild) with CF in GB 93

Table 3.1: Explanation of replacing a missing postcode 113

Table 3.2: Distribution of the number of unique postcodes between 2008 and 2017 (n=527) 113

Table 3.3: Comparison of characteristics of pwCF with a residential postcode in the Greater London Area (GLA) and all pwCF in the UK CF Registry in 2016 121

Table 3.4: Summary of $ppFEV_1$ measurements in pwCF with a residential postcode in the GLA between 2008 and 2017 123

Table 3.5: Summary of yearly, weekly and daily pre-annual review estimated concentration of $PM_{2.5}$ at the residential postcode of pwCF in the GLA between 2008 and 2017 125

Table 3.6: Estimated change in mean $ppFEV_1$ associated with annual exposure of $PM_{2.5}$ among pwCF with a residential postcode in the GLA between 2008 and 2017 (n= 393, $ppFEV_1$ measurements = 3,333) 128

Table 3.7: Summary of the change in AIC with the addition of a covariate in the regression model 130

Table 3.8: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF who did not change their postcode in the GLA between 2008 and 2017 (n= 144, ppFEV$_1$ measurements = 1,225).....132
Table 3.9: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among children pwCF (6 – 15 years) with a residential postcode in the GLA between 2008 and 2017 (n= 122, ppFEV$_1$ measurements = 1,011) 135
Table 3.10: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among adults pwCF (≥16 years) with a residential postcode in the GLA between 2008 and 2017 (n= 271, ppFEV$_1$ measurements = 2,322) 138
Table 3.11: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF with severe genotype (n= 302, ppFEV$_1$ measurements = 2,587) 141
Table 4.1: Number of people with CF (pwCF) living within specific distances from a permitted composting site (PCS) in the UK in 2016 164
Table 4.2: Mean ppFEV$_1$ within and beyond 4 km of a PCS in pwCF in the UK in 2016.....166
Table 4.3: Mean #IVdays within and beyond 4 km of a PCS in pwCF in the UK in 2016.....167
Table 4.4: Linear and Poisson regression analysis of the relationship between ppFEV$_1$ and #IVdays to residential postcode distance from PC in 2016. 169
Table 4.5: Linear regression analysis of the relationship between ppFEV$_1$ and residential postcode distance from PCS among children and adults with CF in the UK in 2016 170
Table 4.6: Respiratory microbiology, fungal disease and mean ppFEV$_1$ among children and adults with CF within and beyond 4km of PCS in the UK in 2016 172
Table 7.1: Linear regressions between air pollutants and lung function in children (age= 6 – 15 years) with CF in the UK CF Registry in GB.....235
Table 7.2: Linear regression between air pollutants (NO$_2$ and PM$_{2.5}$) and lung function in adults (age ≥ 16) with CF in the UK CF Registry in GB who didn't change their residential postcode between 2012 and 2016 236
Table 7.3: Zero-inflated negative binomial regression analysis between NO$_2$ and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB.....237
Table 7.4: Zero-inflated negative binomial regression analysis between NO$_2$ and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB.....238
Table 7.5: Zero-inflated negative binomial regression analysis between PM$_{2.5}$ and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB.....239

Table 7.6: Zero-inflated negative binomial regression analysis between $PM_{2.5}$ and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB.....240

List of Figures

Figure 1.1: Total number of deaths by risk factors in the UK in 2017 34

Figure 1.2: UK annual emissions of NO_2 major sources in 2021 (million tonnes) 36

Figure 1.3: Mean annual concentration in the UK between 1997 and 2021 37

Figure 1.4: UK annual emissions of $PM_{2.5}$ major emission sources in 2021 (thousand tonnes) 39

Figure 1.5: Annual concentrations of $PM_{2.5}$ in the UK, 2009 to 2022 40

Figure 2.1: Conversion of individual postcode to Thiessen polygon boundary 53

Figure 2.2: a: NO_2 concentrations in GB 2015, b: $PM_{2.5}$ concentrations in GB 2015 58

Figure 2.3: Flow diagram of the inclusion and exclusion criteria used for pwCF in GB in the 2016 UK CF Registry database 72

Figure 2.4: Scatter plot of the relationship between $ppFEV_1$ and NO_2 (top) and $PM_{2.5}$ (bottom) in $\mu g/m^3$ in children (left) and adults (right) with CF in GB 78

Figure 2.5: Scatter plot of the relationship between #IVdays and NO_2 and $PM_{2.5}$ in $\mu g/m^3$ in children and adults with CF in GB 79

Figure 2.6: Linear regressions between air pollutants (NO_2 and $PM_{2.5}$) and lung function in children (age 6 – 15 years) with CF in the UK CF Registry in GB 83

Figure 2.7: Linear regression between air pollutants (NO_2 and $PM_{2.5}$) and lung function in adults (age ≥ 16) with CF in the UK CF Registry in GB 85

Figure 2.8: Zero-inflated negative binomial regression analysis between NO_2 and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB 88

Figure 2.9: Zero-inflated negative binomial regression analysis between NO_2 and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB 89

Figure 2.10: ZINB regression analysis between $PM_{2.5}$ and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB 90

Figure 2.11: Zero-inflated negative binomial regression analysis between $PM_{2.5}$ and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB 91

Figure 3.1: Flow diagram of the inclusion and exclusion criteria used for people with CF (pwCF) in the London UK CF Registry 119

Figure 3.2: Overall mean $ppFEV_1$ trend among pwCF with a residential postcode in the GLA between 2008 and 2017 124

Figure 3.3: Mean ppFEV$_1$ and yearly, weekly and daily PM$_{2.5}$ concentration between 2008 and 2017 in the GLA.....126
Figure 3.4: Comparison of the adjusted* estimated change in mean ppFEV$_1$ by exposure of PM$_{2.5}$ among pwCF with a residential postcode in the GLA between 2008 and 2017.....129
Figure 3.5: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF who did not change their residential postcode in the GLA between 2008 and 2017 133
Figure 3.6: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF who did not change their residential postcode in the GLA between 2008 and 2017 134
Figure 3.7: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among children pwCF (6 – 15 years) with a residential postcode in the GLA between 2008 and 2017 136
Figure 3.8: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among children pwCF (6 – 15 years) with a residential postcode in the GLA between 2008 and 2017 137
Figure 3.9: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among adults pwCF (≥16 years) with a residential postcode in the GLA between 2008 and 2017 139
Figure 3.10: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among adults pwCF (≥16 years) with a residential postcode in the GLA between 2008 and 2017 140
Figure 3.11: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF with a severe genotype in the GLA between 2008 and 2017.....142
Figure 3.12: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF with a severe genotype in the GLA between 2008 and 2017 143
Figure 4.1: Flow diagram of the inclusion and exclusion criteria used for people with CF (pwCF) in the 2016 UK CF Registry database.....163
Figure 4.2: Map of the number of pwCF by postcode area and permitted composting sites (PCSs), in the UK in 2016 based on the centroid of postcodes provided by the Environment Agencies of England (n=209), Scotland (n=21), Wales (n=24) and Northern Ireland (n=02) 165

Figure 7.1: Boxplot of the relationship between ppFEV$_1$ and NO$_2$ by covariates in pwCF in GB 219
Figure 7.2: Boxplot of the relationship between ppFEV$_1$ and NO$_2$ by *P. aeruginosa* in pwCF in GB.....220
Figure 7.3: Boxplot of the relationship between ppFEV$_1$ and NO$_2$ by genotype in pwCF in GB 221
Figure 7.4: Boxplot of the relationship between ppFEV$_1$ and PM$_{2.5}$ by covariates in pwCF in GB.....222
Figure 7.5: Boxplot of the relationship between ppFEV$_1$ and PM$_{2.5}$ by *P. aeruginosa* in pwCF in GB.....223
Figure 7.6: Boxplot of the relationship between ppFEV$_1$ and PM$_{2.5}$ by genotype in pwCF in GB.....224
Figure 7.7: Boxplot of the relationship between #IVdays and NO$_2$ by covariates in pwCF in GB.....225
Figure 7.8: Boxplot of the relationship between #IVdays and NO$_2$ by *P. aeruginosa* in pwCF in GB.....226
Figure 7.9: Boxplot of the relationship between #IVdays and NO$_2$ by genotype in pwCF in GB 227
Figure 7.10: Boxplot of the relationship between #IVdays and PM$_{2.5}$ by confounders in pwCF in GB.....228
Figure 7.11: Boxplot of the relationship between #IVdays and PM$_{2.5}$ by *P. aeruginosa* in pwCF in GB.....229
Figure 7.12: Boxplot of the relationship between #IVdays and PM$_{2.5}$ by genotype in pwCF in GB.....230
Figure 7.13: Boxplot of the relationship between subgroup of #IVdays and air pollutants (NO$_2$ and PM$_{2.5}$) in pwCF in GB 231
Figure 7.14: Boxplot of the relationship between subgroup of #IVdays and air pollutants (NO$_2$ and PM$_{2.5}$) in pwCF in GB 232
Figure 7.15 : ppFEV$_1$ by quintile of NO$_2$ Concentrations in GB.....233
Figure 7.16: ppFEV$_1$ by quintile of PM$_{2.5}$ Concentrations in GB 234
Figure 7.17: Correlation between mean ppFEV$_1$ between 2008 and 2017.....241
Figure 7.18: Correlation between annual PM$_{2.5}$ measurement between 2008 and 2017 242

Figure 7.19: Comparison of correlation between ppFEV$_1$ between 2008 and 2017 by age group.....243

Abbreviations

Abbreviation	Explanation
µg/m^3	Micrograms per cubic meter
#IVdays	Number of intravenous days on antibiotics
A. fumigatus	*Aspergillus fumigatus*
ABPA	Allergic bronchopulmonary aspergillosis
AIC	Akaike information criterion
ANOVA	Analysis of Variance
GIS	Geographic Information System.
AURN	Automatic Urban and Rural Network
B. cepacia	*Burkholderia cepacia*
BMI	Body Mass Index (weight/height*2)
CF	Cystic fibrosis
CFFPR	CF Foundation patient registry
CFRD	CF-related diabetes
CFTR	Cystic fibrosis transmembrane conductance regulator
*CFTR*2	Clinical and Functional Translation of CFTR
CHASE	The Child Heart And Health Study in England
CHILA	Children's Hospital Los Angeles
CIs	Confidence intervals
COPD	Chronic obstructive pulmonary disease
CV R^2	Cross-validated coefficient of determination
DEFRA	Department for Environment, Food & Rural Affairs
ECFS	European Cystic Fibrosis Society
EEA	European environment agency
ESRI	Environmental systems research institute, inc.
FEV_1	Forced expiratory volume in one second
FVC	Forced vital capacity
GB	Great Britain
GLA	Greater London Area
GLI	Global Lung Function Initiative
H. influenzae	*Haemophilus Influenzae*
HAPs	Hazardous air pollutants
HEPA	High-efficiency particulate air
IMD	Index of multiple deprivations
IQR	Interquartile range
IRR	Incidence rate ratio
IV	Intravenous
LCI	Lung clearance index
LEZ	Low Emission Zone
LSOA	Lower layer super output area
LUR	Land-use regression
MF	Minimal function
NHS	National health service
NO_2	Nitrogen dioxide
NOx	Nitrogen oxides
O_3	Ozone

ONS	Office for National Statistics
OR	Odds ratio
P. aeruginosa	*Pseudomonas aeruginosa*
PEx	Pulmonary exacerbation
PIAMA	Prevention and Incidence of Asthma and Mite Allergy
PM	Particulate matter
PM_{10}	Particulate matter with a median cut-off diameter of 10µm
$PM_{2.5}$	Particulate matter with a median cut-off diameter of 2.5µm
$ppFEV_1$	Percent predicted forced expiratory volume in one second
pwCF	People with cystic fibrosis
RF	Residual function
S. aureus	*Staphylococcus aureus*
SAHSU	Small Area Health Statistics Unit
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SD	Standard deviation
SOx	Sulfur Oxides
UK	United Kingdom
ULEZ	Ultra-Low Emission Zone
VOCs	Volatile organic compounds
WHO	World Health Organization
ZINB	Zero-inflated negative binomial

Structure of the Project

My PhD thesis is organised into five chapters, with three key chapters for each of the main objectives. By dedicating individual chapters to key areas, the thesis aims to present a focused and in-depth analysis of each research objective.

The first chapter of the thesis serves as an introduction, by providing an overview of the research topic and its significance. It begins with a general introduction to cystic fibrosis (CF), followed by a discussion of various classifications based on genetic mutations and their limitations. Then highlighting the clinical manifestations and progression of the disease. Following the introduction to CF, I discussed the role of environmental factors, with a focus on air pollutants and bioaerosols, on respiratory health and highlighted the need for further investigation in the CF population. Finally, the chapter concludes by outlining the aims and objectives of the research, setting the stage for the subsequent chapters that will delve into each objective in detail.

The three key chapters address the primary objectives of the research and investigate specific aspects of the relationship between air pollutants, bioaerosols, and disease severity in CF. These chapters represent the core of my thesis and contain detailed analyses and findings that contribute to the existing body of knowledge about the role of environmental factors in CF.

In the first key chapter of my research thesis, I focused on exploring the associations between air pollutants (specifically $PM_{2.5}$ and NO_2) and the disease severity in CF in Great Britain. The chapter began with a background on the topic, followed by a detailed description of the methodology employed. I used a semi-individual cross-sectional study design combining high-quality health data from the 2016 UK CF Registry and employed Linear and Poisson regressions to explore the association between air pollutant exposures and disease severity. The results of the analysis were then presented, emphasizing the associations found between air pollutant exposure and disease severity in CF. The discussion section provided a critical analysis of the findings in relation to previous research. Overall, this chapter contributes to our understanding of the impact of $PM_{2.5}$ and NO_2 on disease severity in CF in the specific context of Great Britain.

In my second key chapter, I narrow the focus to the association between $PM_{2.5}$ and Lung function in CF in Greater London, the capital city of Great Britain. The chapter describes the background, study design, methodology and statistical methods used for analysing the longitudinal data. I used a longitudinal study design with repeated measures of annual lung function and linear mixed models to determine the annual rate of change in lung function in CF associated with $PM_{2.5}$ exposure over time (2008 to 2017). The findings are discussed in relation to previous research, providing insights into the specific impact of $PM_{2.5}$ pollution on disease severity in CF in the London population.

In my last key chapter, I describe the relationship between proximity to a composting site in the UK, a proxy for bioaerosol exposure, and disease severity in CF in the UK. The chapter provides an overview of the existing literature on the potential health effects of composting sites and details the methodology. Linear and Poisson regressions were used to explore the health risk of living near composting sites on disease severity in CF in the UK. The results of the study are presented, showcasing the impact of living near a composting site on lung function and pulmonary exacerbations among CF. The findings are discussed in the context of existing knowledge, providing insights into the potential health risks associated with such environmental exposures and recommendations for future research.

The final chapter of my thesis serves as an overall discussion of the key findings presented in the previous chapters. I begin by summarising the main aims of my work and the key findings obtained in three key chapters. The discussion chapter then delves deeper into the implications of these findings, considering their relevance to the UK CF trust and CF community and future research directions, highlighting the need for further investigations and preventive measures.

1 Chapter I: General Introduction

1.1 Cystic Fibrosis

Cystic fibrosis (CF) is an autosomal recessive disorder caused by mutations in the CF transmembrane conductance regulator (*CFTR*) gene, which is located on the long arm of chromosome 7 (Bobadilla *et al.*, 2002). The US Cystic Fibrosis Foundation estimated that 105,000 people have CF across 94 countries in the world, and there are 40,000 pwCF in the US (Cromwell *et al.*, 2023; Cystic Fibrosis Foundation, n.d.). Over 55,000 pwCF live across Europe (39 member countries), while there are over 12000 people with CF (pwCF) in the UK (CF Europe., 2023; UK CF Registry., 2021). It is one of the most commonly inherited life-shortening diseases in people of Caucasian ancestry, with a birth prevalence of one in every 2,500 live births in the UK (da Silveira Fleck *et al.*, 2021; UK CF Registry., 2020). In the 1950s, CF was mainly a disease of children, as most of them did not survive beyond infancy (Elborn, 2016). Modern advances in early diagnosis and treatment methods have led to a considerable improvement in the median life expectancy of pwCF (Coverstone & Ferkol, 2020). The median predicted survival (i.e., the age beyond which half of pwCF born today are expected to live) is now 53 years old, leading to a rising number of CF adults in the UK (UK CF Registry., 2021). This improvement in survival could be attributable to more recent, effective therapies like *CFTR* modulators, regular monitoring and early intervention, advancements in medical knowledge, nutritional support and specialised CF centres (Balfour-Lynn & King, 2022; Keogh *et al.*, 2020). However, pwCF continue to experience substantial morbidity throughout their lives, requiring persistent research efforts to better understand CF pathogenesis and encourage the development of innovative therapies to improve the overall health of pwCF.

1.2 Disease classification

More than 2,000 mutations in *CFTR* have been identified so far (Clinical and Functional Translation of CFTR (CFTR2). 2022b). The homozygous *F508del* mutation is the most common worldwide, accounting for almost 70% of all *CFTR* mutations (Kerem *et al.*, 1989). The five most prevalent *CFTR* mutations in the UK are *F508del* (89%), *R117H* (6.2%), *G551D* (5.7%), *G542X* (3.6%), and *621+1G->T* (2.6%) (UK CF Registry., 2021). *CFTR* gene mutations have a wide spectrum of disease severity. These mutations have been classified in a variety of ways. By classifying these mutations, we increase our knowledge and understanding of underlying

molecular abnormalities, which helps in the diagnosis, guides targeted intervention, and helps to predict disease severity. The first classification which is commonly used is based on three groups: homozygous *F508del*, heterozygous *F508del* and 'other.' In the homozygous *F508del* group, individuals have the *F508del* mutation on both alleles, and the symptoms are often severe in these individuals. Individuals in the heterozygous *F508del* group have *F508del* on one allele and a different *CFTR* mutation on the second allele. These individuals tend to have a milder phenotype than people who have the homozygous *F508del* mutation. Individuals in the "other" group have two *CFTR* mutations other than *F508del*. There is a wide spectrum of mutations that an individual might have; thus, the clinical manifestations and severity of their diseases can vary greatly. This classification is therefore a simplified representation of the spectrum of clinical severity in pwCF and does not fully account for the diversity of clinical phenotypes. Therefore, this classification is often used for a broad classification of *CFTR* and a general understanding of these mutations in clinical and research settings. This has also limitations as it does not fully consider the clinical phenotype, for example, *F508del/G542X* and *F508del/R117H* are both classified as heterozygous *F508del*, but the former is more likely to result in a severe phenotype with pancreatic insufficiency, whereas the latter is more likely to result in a mild phenotype with pancreatic sufficiency.

Another commonly used classification is a five (or six) group classification based on protein synthesis and function (De Boeck & Amaral, 2016): Class I mutations produce no protein; Class II mutations have a protein processing deficiency; Class III mutations include all gating mutations; Class IV mutations have conduction defects; and Class V mutations produce an insufficient quantity of protein. A further group (Class VI) is considered by some, which leads to unstable *CFTR* protein and increased turnover. However, this classification system has significant limitations as there is considerable overlap between the groups and not all mutations have been studied sufficiently to be able to classify them (Haq *et al.*, 2022; Veit *et al.*, 2016).

1.3 Clinical manifestations and disease progression

Although CF is a systemic disease that also affects the pancreas, the digestive system, and infertility, the reparatory system remains the main contributor to the severe morbidity and

mortality in pwCF (Blayac *et al.*, 2022; Elborn, 2016). Therefore, this thesis focuses on respiratory disease in pwCF.

CFTR protein functions predominantly as a chloride channel in epithelial cells, regulating the movement of salt and water in and out of cells, resulting in clinical complications particularly affecting the lungs and the gastrointestinal system. Clinical respiratory manifestations such as chronic airway blockage, recurrent respiratory infections, and progressive lung damage are key characteristics of CF disease. pwCF produces thick secretions, which provide a suitable environment for the growth of pathogens, including bacteria and fungi. This leads to recurrent respiratory infections called 'pulmonary exacerbations' (PEx) and an inflammatory environment resulting in progressive respiratory disease (Ranganathan *et al.*, 2013).

PEx are characterised by an acute exacerbation of respiratory symptoms, such as increased cough, sputum production, and breathing difficulty, and are a significant cause of morbidity and mortality in pwCF (Stanford *et al.*, 2021). PEx are associated with a decline in lung function, disease progression, increased healthcare cost, and a poor quality of life in pwCF (Don & Christopher Hooper, 2013; Nelson, 2020). The frequency and severity of PEx in pwCF are influenced by a wide range of factors, including age, sex, *CFTR* genotype, bacterial and viral infection, BMI, comorbidities, socioeconomic status, and environmental factors (Nelson, 2020). Early recognition and treatment of PEx are important for preventing permanent lung damage and improving long-term outcomes in pw CF. Recurrent PEx and persistent inflammation are associated with reduced lung function and poor survival in pwCF. A decline in lung function has been associated with significant morbidity and mortality in pwCF (Breuer *et al.*, 2018; Harun *et al.*, 2016). Lung function is a recognised indicator of disease progression in CF and has been used in numerous studies to assess the effectiveness of treatment and the disease's progression in pwCF (Caley *et al.*, 2021; Szczesniak *et al.*, 2017a; Taylor-Robinson *et al.*, 2012).

1.4 Infections in pwCF

Respiratory infections are the leading cause of morbidity and mortality among the CF population. Repeated bacterial respiratory infections are the main cause of this high morbidity and mortality in pwCF (Ibrahim & Abdul, 2017; Kiedrowski & Bomberger, 2018). Numerous bacterial species, such as *Pseudomonas aeruginosa (P. aeruginosa)*,

Staphylococcus aureus (S. aureus) and *Haemophilus influenzae (H. influenzae)*, have been found in the lungs of pwCF but *P. aeruginosa* and *S. aureus* are the most prevalent bacterial species detected in pwCF lungs (Fischer *et al.*, 2021). These bacteria are associated with an increase in PEx and a rapid decline in lung function in pwCF (Garcia-Clemente *et al.*, 2020; Junge *et al.*, 2016). Substantial work has been done on the treatment and prevention of bacterial infection among pwCF (Fischer *et al.*, 2021; Gilligan, 2014; Hatziagorou *et al.*, 2020; Lyczak *et al.*, 2002). This led to substantial increases in the life expectancy of pwCF but also resulted in an increased risk for the colonisation of fungi.

Several fungal species, including *Aspergillus fumigatus (A. fumigatus)*, Candida species, *Scedosporium* species, and non-tuberculous mycobacteria, are common fungi that present in the lungs of pwCF, with *A. fumigatus* being the most prevalent (Magee *et al.*, 2021; Williams *et al.*, 2016). Allergic bronchopulmonary aspergillosis (ABPA) is the most well-known fungal disease in pwCF, with a reported prevalence of between 2 and 25% (Maturu & Agarwal, 2015). Fungal infections may be associated with a decline in lung function and PEx (Amin *et al.*, 2010; Saunders *et al.*, 2016). There has been a steady increase in the number of fungal infections among pwCF since 2007 (UK CF Registry., 2020). *A. fumigatus*, the most prevalent fungus, cases were only 6% in 2007 and increased to 15% in 2016 (Delfino *et al.*, 2019; UK CF Registry., 2020). *Aspergillus spp.* isolation from sputum cultures in pwCF increases with age and may reach 46–78% in CF adults (Bakare *et al.*, 2003). The mechanisms underlying this trend remain unclear, but existing evidence points towards a substantial contribution from environmental factors (Kanchongkittiphon *et al.*, 2015; Kuhn & Ghannoum, 2003; Mendell *et al.*, 2011).

1.5 Environmental factors

There is significant variation in CF lung disease, even in pwCF with the same *CFTR* genotype, and nearly half of this variation is attributed to non-genetic factors (Collaco *et al.*, 2010; Drumm *et al.*, 2012). It is likely that environmental factors such as air quality (Blayac *et al.*, 2022; Brugha *et al.*, 2018), tobacco smoke (Oates *et al.*, 2020), season and climate (Qvist *et al.*, 2019; Ramsay *et al.*, 2016), housing environment and socioeconomic status (Oates *et al.*, 2020; Szczesniak *et al.*, 2020; Taylor-Robinson *et al.*, 2014) can influence disease severity in CF. Recently, my colleague Olivia Leigh Murrin conducted a systematic study for her PhD work on the association between environmental variables and CF. Her research emphasises several

environmental factors in CF, including Living in an urban area is recognised as a possible environmental risk factor in pwCF. Moreover, despite improved healthcare for individuals with CF, the debate on whether bacterial and fungal infections are mostly transmitted by person-to-person contacts, environmental factors, or both is still ongoing. The role of environmental factors in the severity of CF has been largely neglected so far, particularly in the UK. Studies conducted in the US and continental Europe have shown: i) an association between ambient temperature and lung function (Collaco *et al.*, 2011); ii) an association between air pollutants (NO_2, PM_{10}, and ambient concentrations of ozone) and pulmonary exacerbations (Goeminne *et al.*, 2013) and iii) association between dampness/moulds and lung function (Norbäck *et al.*, 2011). There was also a strong association (based on ORs) between air pollution, antibiotic use and ambient temperature (Goeminne *et al.*, 2013).

1.5.1 Air pollution

Air pollution is a major environmental and public health problem, and it is affecting everyone across the globe. Air pollution is the fourth highest risk factor for human health in the world (Rafaj *et al.*, 2018). The World Health Organisation (WHO) estimates that more than 99% of the world's population lives in areas with poor air quality and air pollution levels that are much higher than WHO regulatory limits (10 μg/m^3) (WHO., 2022). Each year, air pollution is attributed to more than seven million premature deaths and more than 100 million DALYs (disability-adjusted life years) worldwide (Cohen *et al.*, 2017; WHO., 2022). Outdoor air pollution is one of the major global hazards to human health; there were over 4 million deaths attributed to outdoor air pollution in 2019 (WHO., 2022).

Every year in the UK, air pollution is responsible for between 28,000 and 36,000 deaths, and the total cost of air pollution to the National Health Service (NHS) is anticipated to be £1.6 billion between the years 2017 and 2025 (Office for Health Improvement & Disparities., 2022). Given the significant health burden and detrimental impacts of air pollution on human health, research studies are imperative to better understand their complex effects and develop effective strategies to address this global issue.

Air pollution is a complex mixture of particles and gases and comes from a variety of sources, such as emissions from industries, exhaust from automobiles, power plants, agricultural activities, and fuel burning in homes. These sources may produce a range of pollutants,

including particulate matter (PM), nitrogen oxides (NOx), ozone (O_3) sulfur oxides (SOx), volatile organic compounds (VOCs), and hazardous air pollutants (HAPs). Each of these pollutants has a distinct source and diverse characteristics.

1.5.2 Impact of air pollution on health

Air pollution has been associated with a higher rate of morbidity and mortality reported in previous studies (Di *et al.*, 2017; Lelieveld *et al.*, 2015; Schraufnagel *et al.*, 2019a; Schraufnagel *et al.*, 2019b). Air pollution has been associated with cancer (Balmain, 2023; Gourd, 2022), heart disease (Rajagopalan & Landrigan, 2021), stroke (Verhoeven *et al.*, 2021) and respiratory diseases (Monoson *et al.*, 2023). In 2017, 6% of worldwide deaths were attributed to outdoor air pollution Traffic emissions are one of the major contributors to air pollution, therefore, traffic-related air pollutant concentrations vary on major roads (Baldwin *et al.*, 2015).

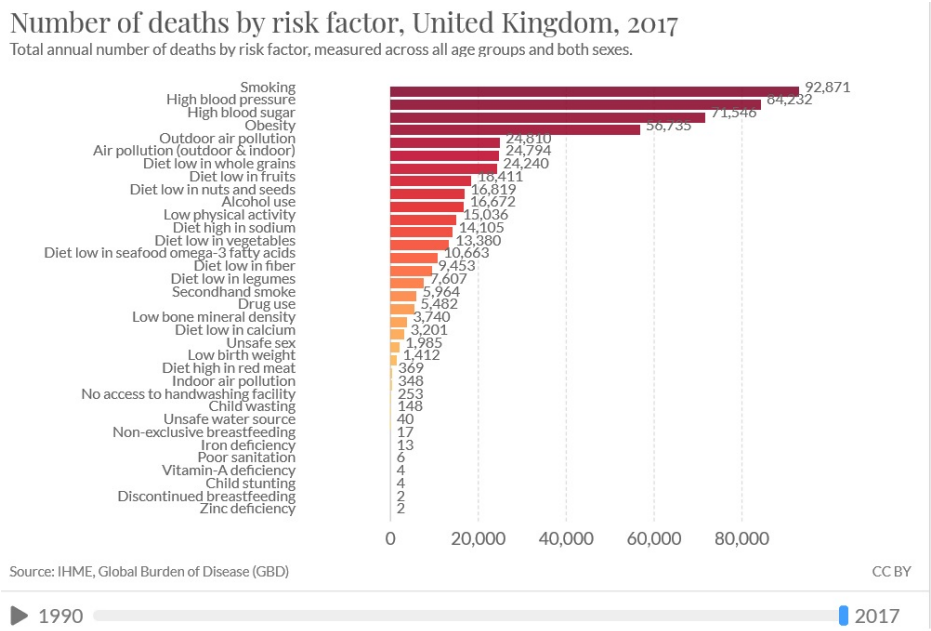


Figure 1.1: Total number of deaths by risk factors in the UK in 2017 (Hannah Ritchie. & Max Roser., 2022).

As shown in ***Figure 1.1,*** air pollution is one of the top four leading risk factors for death in the UK in 2017(Hannah Ritchie. & Max Roser., 2022). Air pollution is associated with poorer health outcomes among older adults than young adults (Wong *et al.*, 1999). Of air pollutants, NO_2 and $PM_{2.5}$ are associated with worse health effects compared to exposure to other pollutants

(Burnett *et al.*, 2018; Hanigan *et al.*, 2019; Schraufnagel *et al.*, 2019b). In my thesis, I focused on these two pollutants, as they are the most prevalent, reflect the main sources of emission, extensively studied air pollutants and are known for having detrimental effects on respiratory health in a variety of populations (Liang *et al.*, 2023; Wang *et al.*, 2022a). Also, there is detailed data from monitoring stations and a complex model available for these air pollutants to assess human exposure at a fine scale.

The other reason for focusing on these two air contaminants is that they are more prevalent in urban areas due to vehicle emissions, industrial activity, and power plants (Cooper *et al.*, 2019). To my knowledge, there is no published data tracking air pollution exposure for pwCF. However, trends in overall air pollution, including $PM_{2.5}$ and NO_2, have shown a decline between 2009 and 2022 within the general population of the UK (DEFRA., 2023b; DEFRA., 2023c). The implementation of London's Low Emission Zone (LEZ) in 2008, followed by the introduction of the Ultra-Low Emission Zone (ULEZ) in 2019, has also played a significant role in improving air quality in London (Eleonora Fichera. *et al.*, 2023). These policy initiatives have resulted in substantial reductions in vehicle emissions, notably NOx by 26% and PM by 19% (Greater London Authority., 2024). The subsequent decline in harmful emissions has translated into significant improvements in public health outcomes over a relatively short period (Chamberlain *et al.*, 2023). Nevertheless, further research is warranted to comprehensively assess the long-term health impacts of these measures. It is likely that air pollution exposure for pwCF in the UK has decreased, but more specific data is needed. Understanding the specific impact of these air pollutants on disease severity in CF can provide valuable insights into the potential environmental risk factors contributing to disease severity. Studying the relationship between NO_2 and $PM_{2.5}$ and disease severity in pwCF may help to better understand the disease prognosis and provide insights for targeted treatments and strategies to improve the respiratory health and quality of life of pwCF.

1.5.3 Nitrogen dioxide (NO_2)

NO_2 is a man-made, reddish-brown, highly reactive gas that is water-soluble and a strong oxidant. It is one of the most harmful air pollutants, with significant implications for the environment and public health. There are several different sources of NO_2 emissions in the UK, with road transport being the main contributor (almost 70%). In 2021, 0.184 million

tonnes of emissions were attributed to road transport (***Figure 1.2***) (DEFRA., 2023a). NO_2 is a good proxy for measuring air pollution related to traffic and transport. Vehicle emissions, especially those from diesel vehicles, release large amounts of NO_2 into the environment. Industrial processes, which include manufacturing, chemical production, and waste incineration, are another significant source of NO_2 emissions. It contributed 0.135 million tonnes of NO_2 emissions in 2021 (DEFRA., 2023a). Power plants, which are used to generate energy, are also substantial sources of NO_2 emissions (0.128 million tonnes in 2021) (DEFRA., 2023a). Domestic heating and cooking are other sources of NO_2 emissions. Using natural gas, coal and wood in households can release NO_2 into the environment. Agricultural activities, such as fertiliser and livestock farming, are also sources of NO_2 emissions.

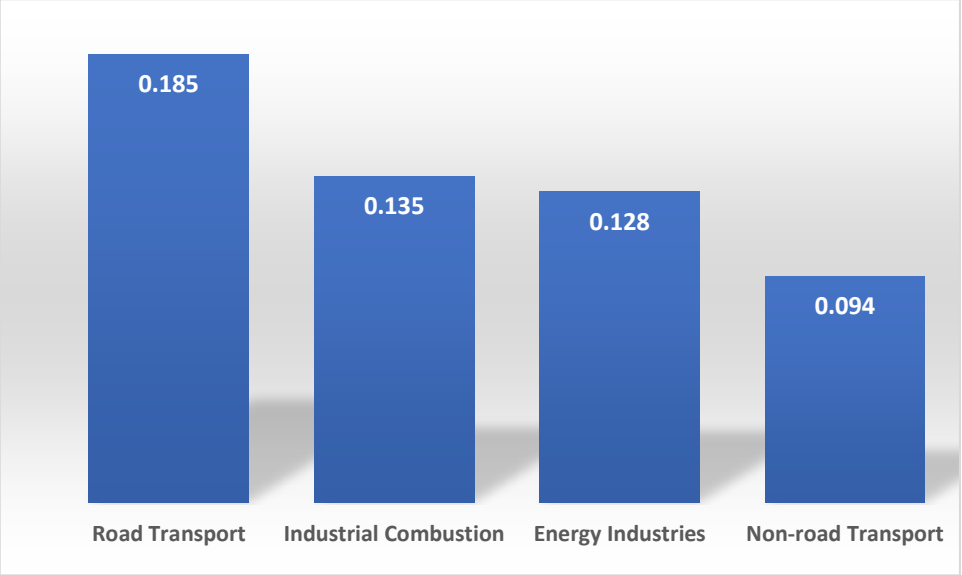


Figure 1.2: UK annual emissions of NO_2 major sources in 2021 (million tonnes)

The UK Air Quality Standards 2010 have defined NO_2 concentration limits to protect the public's health. According to the Air Quality Standards Regulations 2010, the legal limit for NO_2 is 40 µg/m^3 and exposure to hourly mean NO_2 should not exceeded200 µg/m^3 more than 18 times in a year (DEFRA., 2023c).

According to the Department for Environment, Food, & Rural Affairs, the mean annual NO_2 concentration in the UK has generally been on the decline between 1970 and 2021. In urban areas, NO_2 concentrations decreased to 19.6 µg/m^3 in 2019 from 46.0 µg/m^3 in 1997. Rural

areas in the UK had lower NO_2 concentrations of 17.7 µg/m^3 and 7.2 µg/m^3 in 1997 and 2019, respectively (***Figure 1.3***) (DEFRA., 2023c). The mean annual NO_2 concentration in 2016 was 23.3 µg/m^3 in urban areas compared to 8.7 µg/m^3 in rural areas. The annual mean NO_2 concentration at the roadside reached its lowest point in the time record in 2020, having decreased by 8.21 µg/m^3 (26%) over the previous year. One possible explanation for this rather big decline could be the traffic reduction brought about by COVID-19 restrictions. In 2021, there was an 8% increase in the annual mean NO_2 concentration at the roadside, amounting to 1.8 µg/m^3. This is likely a result of increased road traffic following the lifting of lockdown restrictions. Concentrations in 2022 fell somewhat from 2021 and remained 24% below 2019 levels.

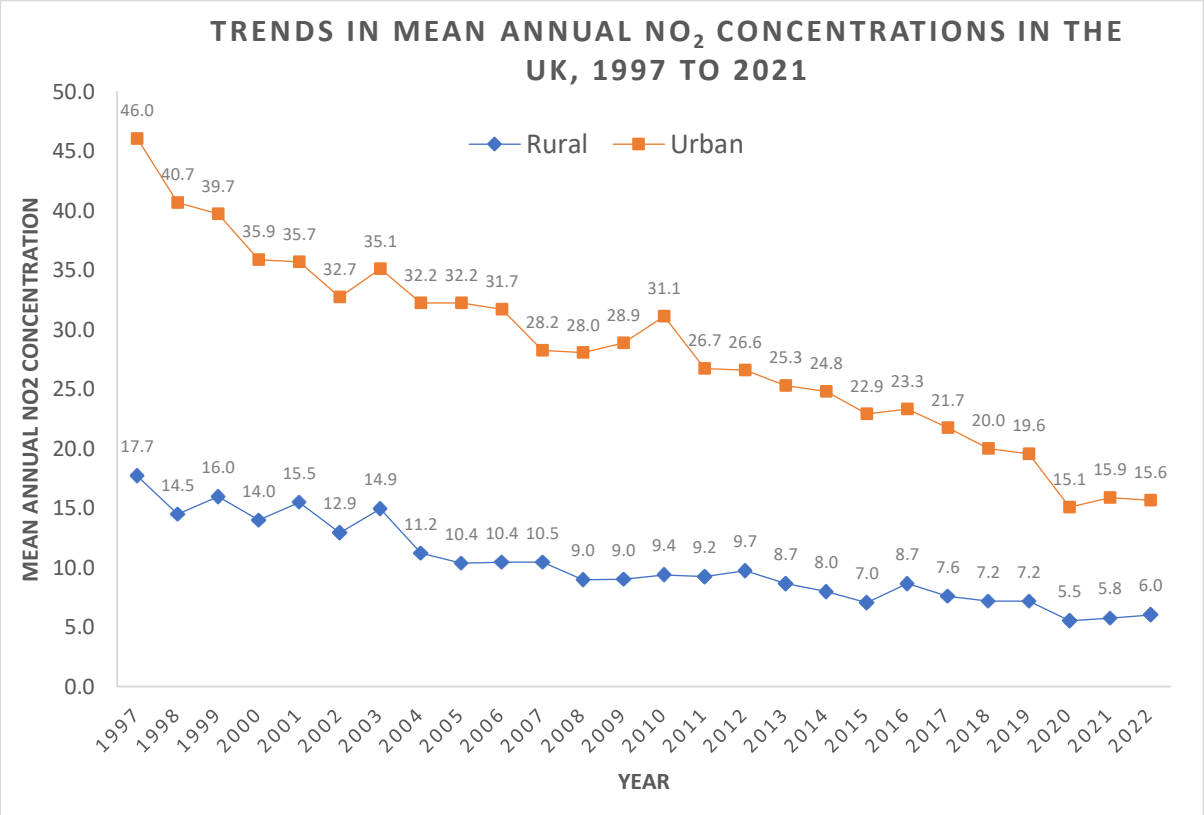


Figure 1.3: Mean annual concentration in the UK between 1997 and 2021
(DEFRA., 2023c)

Exposure to high levels of NO_2 concentrations has been reported to be associated with heart diseases (Gan *et al.*, 2012; Shin *et al.*, 2020), diabetes (Shin *et al.*, 2020), respiratory diseases (Beelen *et al.*, 2008; Chen *et al.*, 2007; Hoek *et al.*, 2013) and neurologic diseases (Wellenius *et al.*, 2005). Poor lung function in adults (Bowatte *et al.*, 2017), and among children (Avol *et*

al., 2001) is also associated with NO_2 concentrations. NO_2 was also found to be associated with various infections, including otitis media (Brauer *et al.*, 2006), eczema (Morgenstern *et al.*, 2008) and ear, nose and throat infections (Brauer *et al.*, 2007). NO_2 also irritates the airways of the lungs and continuous exposure could lead to premature death.

Children and people with respiratory diseases such as chronic obstructive pulmonary disease (COPD), asthma, or CF are more susceptible to the harmful effects of exposure to NO_2, which can have a significant impact on their health. Individuals with pre-existing respiratory conditions have reduced lung function, making them more susceptible to NO_2-induced exacerbations and respiratory symptoms. Research shows that even low levels of exposure to NO_2 can worsen their symptoms and may cause lung function to decline significantly in these vulnerable populations (Tiotiu *et al.*, 2020).

1.5.4 Particulate matter (PM)

PM is a mixture of various particles, including solids and liquids, suspended in the atmosphere and of various sizes, compositions, and origins. PM is generally divided into three categories: coarse particles (PM_{10}) with a diameter of 2.5–10 microns, fine particles ($PM_{2.5}$) with a diameter of 2.5 m or less, and ultrafine fine particles ($PM_{0.1}$) with a diameter of less than 0.1 m.

Like NO_2, PM is emitted by various sources, including domestic heating, especially wood stoves, industrial combustion, industrial processes, road transport and agricultural practices ***(Figure 1.4)***. About 22.70 thousand tonnes of $PM_{2.5}$ emissions in the UK in 2021 were attributed to domestic heating (DEFRA., 2023b). There were 21.60 thousand tonnes of $PM_{2.5}$ emissions from industrial combustion in the UK (DEFRA., 2023b). Whereas industrial processes, such as those associated with manufacturing and power production, contributed 11.18 thousand tonnes of $PM_{2.5}$ emissions in the UK in 2021 and road transport contributed 10.73 thousand tonnes (DEFRA., 2023b). Agriculture-related activities, including the farming of livestock, burning crops, and using pesticides and fertilisers, can all leave behind organic particles, dust, and chemical residues that contribute to $PM_{2.5}$ emissions. Understanding these various sources and their potential contributions is essential for developing effective strategies and policies to reduce $PM_{2.5}$ emissions.

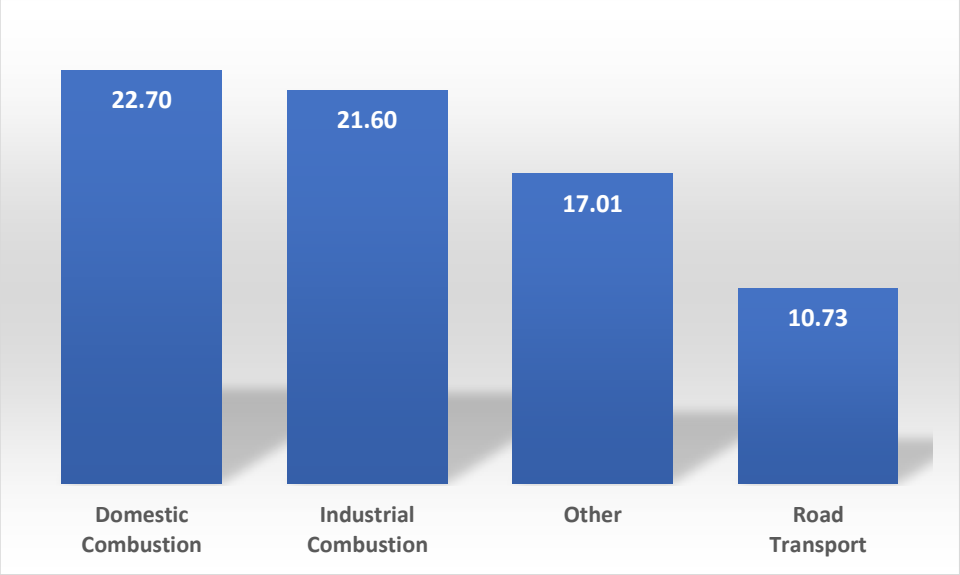


Figure 1.4: UK annual emissions of $PM_{2.5}$ major emission sources in 2021 (thousand tonnes)

Overall, $PM_{2.5}$ concentrations increased from 2009 to 2011 as seen in Figure 1.5. After that, they started to decline towards 2015 and were generally stable until 2019. Then, in 2020, concentrations were lower, and they have stayed comparatively steady ever since. The mean concentrations in 2022 increased by 5% over 2021 levels. The annual average concentrations of $PM_{2.5}$ at urban background sites have shown a decline, decreasing from 12.4 µg/m³ in 2009 to 8.3 µg/m³ in 2022. Similarly, the annual mean concentrations of $PM_{2.5}$ in roadside areas have also exhibited a downward trend, decreasing from 12.8 µg/m³ in 2009 to 8.7 µg/m³ in 2022. Notably, there has been a modest increase of 0.7 µg/m³ since reaching the lowest point in the time series in 2020 (DEFRA., 2023b).

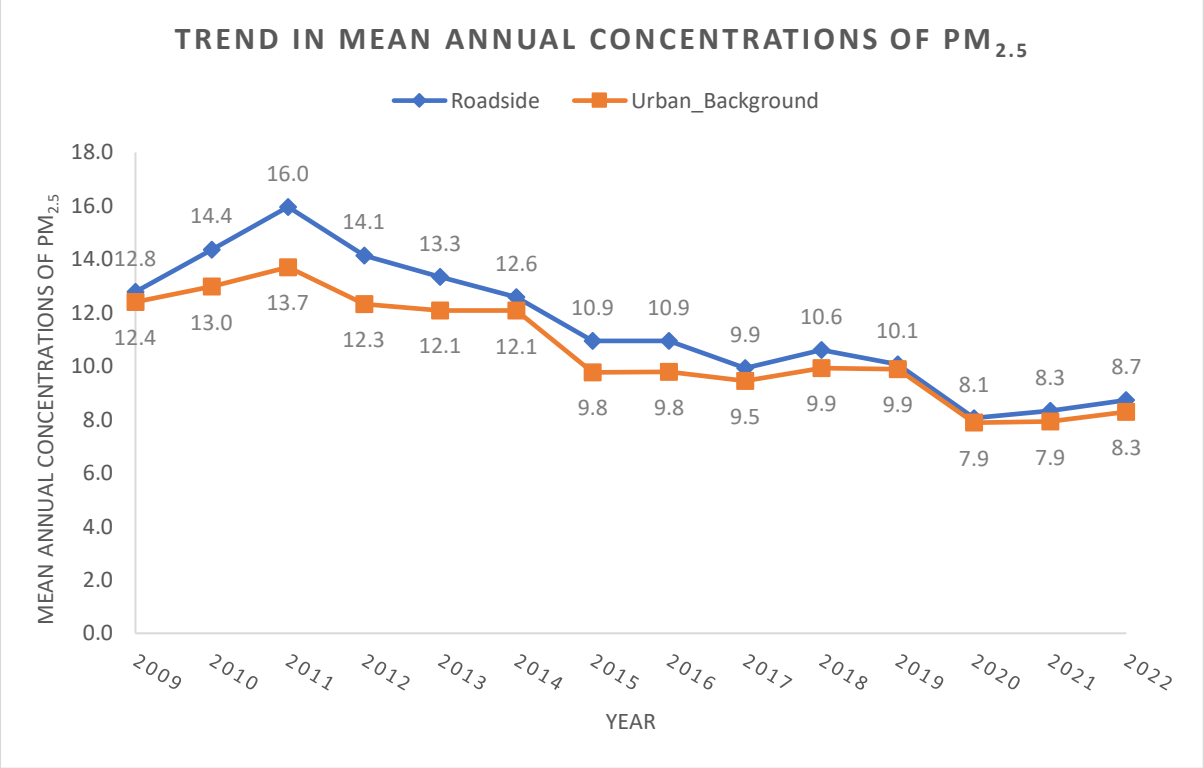


Figure 1.5: Annual concentrations of $PM_{2.5}$ in the UK, 2009 to 2022

A growing body of epidemiological research suggests that exposure to ambient $PM_{2.5}$ significantly increases the risk of respiratory mortality and morbidity. Governments and health organisations around the world are increasingly worried about the detrimental effects of PM on human health (The Lancet, 2006). Globally, 1,702,150 deaths and 47.5 million DALYs were attributed to exposure to ambient $PM_{2.5}$ (Sang *et al.*, 2022). Due to the detrimental impact of $PM_{2.5}$ on human health, WHO set legal mean limits for $PM_{2.5}$ at 10 µg/m^3, but this has been updated to 5 µg/m^3 since there was a greater risk of death associated with $PM_{2.5}$ in areas where the $PM_{2.5}$ was below 10µg/m^3, as reported in multiple meta-analysis studies (Chen & Hoek, 2020; DEFRA., 2021; Orellano *et al.*, 2020). In the UK, the current legal limit for $PM_{2.5}$ is 20 µg/m^3 under the 2010 Air Quality Standards (DEFRA., 2023c). However, recently, the UK government announced a new target for $PM_{2.5}$, aiming for 10 µg/m^3 by 2040 (DEFRA., 2021).

The annual average concentrations of $PM_{2.5}$ at urban background locations declined from 12.4 µg/m^3 in 2009 to 8.3 µg/m^3 in 2022, according to figures from the UK government (DEFRA., 2023b). In 2016, annual average concentrations of $PM_{2.5}$ were 15.6 µg/m^3.

An extensive body of scientific evidence shows that long- and short-term exposures to $PM_{2.5}$, can cause premature death (Bowe *et al.*, 2019) and harmful effects on the cardiovascular system (Hayes *et al.*, 2020), including increased hospital admissions and emergency department visits for heart attacks and strokes (Chang *et al.*, 2022). Scientific evidence also links $PM_{2.5}$ to harmful respiratory effects (Manisalidis *et al.*, 2020; Xing *et al.*, 2016), lung cancer (Coleman *et al.*, 2020), type 2 diabetes (He *et al.*, 2017) and preterm birth (Trasande *et al.*, 2016; Yang *et al.*, 2021b).

$PM_{2.5}$ has a significant impact on vulnerable groups, such as children, older people, individuals with pre-existing health conditions such as CF and individuals with socioeconomic disparities. Children are more vulnerable to the impacts of $PM_{2.5}$ because of their developing lungs and higher inflammatory susceptibility. Similarly, older people are more vulnerable to the higher impact of $PM_{2.5}$ due to age-related physiological changes in their bodies and the comorbidities they have. Furthermore, individuals with pre-existing health conditions are more susceptible to the harmful effects of $PM_{2.5}$ due to compromised health conditions. These inequalities make certain populations more susceptible to the negative health impacts of $PM_{2.5}$ exposure. Understanding the effects of $PM_{2.5}$ on vulnerable groups such as CF is essential for developing tailored interventions to reduce $PM_{2.5}$ exposure and improve their overall well-being.

The effects of NO_2 and $PM_{2.5}$ on pwCF in the UK have not been studied. By examining how these air pollutants increase the severity of the condition, we can better understand the consequences of these air pollutants exposure, find strategies to reduce these risks and improve the health outcomes for pwCF in the UK.

1.5.5 Existing evidence on the relationship between outdoor air pollution and outcomes in pwCF

My research on the impact of air pollution on the CF course identified two published systematic reviews on this topic. First Brugga *et al.* comprehensively assessed past research, establishing major outdoor air pollution sources and critically evaluating air pollution exposure assignment methods (Brugha *et al.*, 2018). They further provided a synthesis of existing evidence on outdoor air pollution's effects on CF, highlighting key findings such as increased pulmonary exacerbation risk and decreased lung function in pwCF associated with air pollution exposure. Building upon this foundation, Blayac *et al.* focused on the mechanistic

understanding of air pollution's role in CF pathogenesis (Blayac *et al.*, 2022). Their review explored cellular and biological responses to both acute and chronic air pollution exposure, incorporating findings from diverse observational and experimental studies, including *in vivo* and *in vitro* studies. In addition, my colleague Olivia Leigh Murrin conducted a systematic review of associations between environmental factors and CF in her PhD dissertation submitted in August 2022 (Olivia Murrin., 2022). Her unpublished review highlighted the complex interplay between environmental factors and CF outcomes, with regional variations playing a significant role. Areas experiencing extreme temperatures and deprivation seemed to have worse CF health, likely due to increased infections and inadequate housing. Climate influenced pulmonary manifestations via increased infections, with both temperature and humidity influencing pathogen survival and transmission. Air pollution, particularly $PM_{2.5}$ and PM_{10}, showed clear links to acute exacerbations, likely through immune system effects and increased infection risk. Long-term exposure was associated with lung function decline in CF patients, emphasizing the need for mitigation strategies. In support of the conclusions drawn in these reviews, it is worth noting that only four studies have investigated the impact of outdoor air pollution on disease severity among pwCF (Farhat *et al.*, 2013; Goeminne *et al.*, 2013; Goss *et al.*, 2004; Jassal *et al.*, 2013). I will briefly summarize these key studies with their strengths and limitations here.

A pivotal cohort study, conducted by Goss *et al.* in 2004, particularly stands out as the most extensive investigation within the pwCF population to date (Goss *et al.*, 2004). The study included 11,484 CF patients over the age of 6 years enrolled in the Cystic Fibrosis Foundation National Patient Registry between 1999 and 2000. Exposure assessment was based on linking mean annual values for air pollutants from the monitoring station within 30 miles of the pwCF's home zip code in the Aerometric Information Retrieval System (AIRS). They reported a 21% increase in the odds of experiencing two or more exacerbations, and a decrease of 24 mL in FEV_1 for every 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure (95% CI: 7-33%). However, no distinct association was observed with other pollutants (NO_2, O_3, CO, or SO_2). The study's strength lies in its large sample size and comprehensive evaluation of air pollution's impact on disease severity. However, a limitation is the reliance on using the nearest monitoring station data for exposure assessment, which may lack precision compared to high-resolution continuous modelled surfaces derived from land-use regression or dispersion models.

Additionally, the study acknowledged the absence of data on tobacco use, climate, weather patterns, and regional variations in medical care, which could have influenced the study's findings.

The study, conducted by Goeminne and colleagues investigated the effect of acute exposure to daily fluctuations in air pollutants (PM_{10}, O_3, and NO_2) on exacerbations among 215 pwCF in a retrospective case-crossover study spanning 12 years (1998–2010) in Belgium (Goeminne *et al.*, 2013). Air pollution exposures for PM_{10}, NO_2, and ozone were estimated using modelling techniques to provide data for 4 km grid squares two days prior, one day prior, and on the day of a pulmonary exacerbation (PEx) onset. A PEx was defined as the use of oral antibiotics either at home or in the hospital, in addition to the use of intravenous (IV) antibiotics at home or in the hospital. This definition differs from that of Goss *et al.,* who only considered the use of IV antibiotics at home or in the hospital for defining a pulmonary exacerbation. The findings revealed significant associations between all three air pollutants and same-day exacerbations. The authors recognised the exclusion of CF participants without exacerbations during the study period, highlighting that the lack of reported exacerbations could also be attributable to missing data for those pwCF. The study's strength lies in its case-crossover design, which employs each pwCF as their own control, effectively minimising the influence of potential confounding factors such as socioeconomic status.

Another study of relatively small size was conducted by Farhat *et al.* to examine the association between short-term fluctuations in air pollutant concentrations and respiratory exacerbations in children and adolescents with CF residing in the highly polluted metropolis of São Paulo, Brazil (Farhat *et al.*, 2013). The study involved 103 participants (median age, 8.9 years) followed for 12 months. Exacerbations were defined as the presence of at least three of the following symptoms: fever, increased sputum production or cough intensity, change in sputum colour, worsening dyspnea, loss of appetite, >10% decrease in FEV_1, or weight loss. Daily mean concentrations of ozone (O_3), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), particulate matter 10 (PM_{10}), and carbon monoxide (CO) were measured from the monitoring stations across the city and compared to the occurrence of exacerbations. A 45.62 $\mu g/m^3$ interquartile range increase in O_3 concentration was associated with a statistically significant 86% increase in the relative risk of respiratory exacerbation two days after exposure (RR =

1.86; 95% CI, 1.14–3.02). Notably, no other investigated pollutants exhibited a significant association with exacerbation occurrence in pwCF. However, methodological limitations, such as unclear clinical data and a less stringent exacerbation definition, raise questions about the role of other pollutants like PM_{10} and NO_2. While both this study and another by Goeminne *et al.* support an association between air pollution and CF exacerbations, discrepancies in findings highlight the need for further research with improved methods, standardized definitions, and consideration of CF patient characteristics to definitively establish the specific air pollutants and mechanisms impacting CF disease progression. Ultimately, understanding the air pollution-CF relationship is crucial for developing strategies to protect affected individuals and improve their quality of life.

Finally, Jassal *et al.* (2013) investigated the association between air pollution (O_3, $PM_{2.5}$) and residential proximity to roadways as a proxy for long-term traffic pollution, given elevated particle concentrations near roads due to vehicle emissions and pulmonary exacerbations in 145 CF patients (Jassal *et al.*, 2013). The 5-year retrospective study utilised the same exacerbation definition as Goss *et al.,* classifying patients into groups with zero or one exacerbation versus two or more exacerbations. While Jassal *et al.* did not find significant differences in O_3 and $PM_{2.5}$ levels between the two exacerbation groups, residential proximity to a roadway was higher for individuals experiencing more than two exacerbations. The study failed to replicate the effects observed by Goss *et al.*, potentially due to low statistical power. Nevertheless, Jassal's study was the first to highlight the influence of traffic proximity on the clinical condition of CF patients, introducing a novel perspective beyond traditional pollutant measures. The major limitations of this study were a small sample size, a limited exposure assessment (using proximity to roadways as a proxy for traffic pollution), and a lack of potential confounding factors (e.g., socioeconomic status, smoking), which Jassal *et al.* also acknowledged. In summary, the research of Jassal *et al.* underscores the relevance of residential proximity to roadways as a factor influencing CF exacerbation frequency, offering valuable insights beyond conventional pollutant measures. This study contributes to the growing body of literature associating outdoor air pollution with CF exacerbations despite methodological limitations in exposure classification. These findings have implications for future predictive models of the clinical outcomes of CF.

1.5.6 Bioaerosols

I considered bioaerosol exposure along with NO_2 and $PM_{2.5}$ air pollutants in my thesis for a number of reasons. First, airborne particles, also known as bioaerosols, are a group of airborne microorganisms that are present in the air together with PMs and come in different shapes and sizes (Caruana, 2011). Exposure to bioaerosols, which can make up 15 to 25% of PM by mass and may contain live or dead bacteria, fungi, viruses, dust, and pollen, has been associated with a number of respiratory health issues, including CF (Roca-Barcelo *et al.*, 2020; Samake *et al.*, 2017). Second, bioaerosols are a different type of emissions, with a more defined point source similar to incinerators or nuclear power plants (British society for ecological medicine., 2018; Desbiolles *et al.*, 2018; Douglas *et al.*, 2017a; Kim *et al.*, 2016). Desbiolles *et al.* found an increased risk of bladder cancer in both men (RR = 1.08, 95%CI: 1.00; 1.17) and women (RR = 1.19; 95%CI: 1.02; 1.39) in a 20-20-km zone around a nuclear power plant in France (Desbiolles *et al.*, 2018). Third, one of the major sources of bioaerosols are composting sites and exposure to these may have a significant negative impact on one's respiratory health. Bioaerosol exposure from composting is a growing concern in the United Kingdom (UK), as it is associated with health risks for individuals living near composting sites, particularly in relation to fungal infections. It is imperative to investigate the effects of living near a composting site on the health of pwCF since there is a significant increase in the prevalence of fungal infections among pwCF and *Aspergillus fumigatus (A. fumigatus)*, a pathogenic fungus that may cause inflammatory, immune, and allergic reactions when inhaled, is a well-known and common component of bioaerosols (Pearson *et al.*, 2015; Poore *et al.*, 2021; UK CF Registry., 2020). Bioaerosols are more likely to affect pwCF compared to the general population, hence I focused on pwCF in this study.

Furthermore, I added a bioaerosol exposure section to my thesis as a result of an opportunity presented in the early stages of my PhD, when I had the pleasure of collaborating with a senior PhD student who had prior experience investigating composting sites in the UK. By adding the bioaerosol exposure dimension, I hope to contribute to the understanding of the potential associations between bioaerosol exposure and disease severity in pwCF.

Despite the limited data on the trend of bioaerosol exposure specifically for pwCF near composting sites in the UK, concerns remain due to the rapid expansion of such facilities.

Between 2010 and 2017, the number of large-scale composting plants in England more than doubled, from 140 to over 300, potentially leading to increased bioaerosol emissions (Robertson *et al.*, 2019). Also, the health effects of these emissions on surrounding communities, particularly pwCF, remain uncertain. Although some studies have observed increased respiratory symptoms and lung function decline in composting workers (Bunger *et al.*, 2000; Pearson *et al.*, 2015; Robertson *et al.*, 2019; van Kampen *et al.*, 2016), the evidence linking objective bioaerosol exposure with health outcomes in this population is scarce. Further research is crucial to refine risk assessment and inform relevant public health recommendations for the CF population living near composting sites.

1.6 Hypothesis

- Higher levels of NO_2 and $PM_{2.5}$ exposure are associated with increased disease severity, as indicated by increased PEx and decreased lung function, in children and adults with CF in GB.
- Higher long-term exposure to $PM_{2.5}$ is associated with a faster decline in lung function compared to a low level of exposure to $PM_{2.5}$ among pwCF residing in Greater London over the study period.
- Living close to a permitted composting site is associated with increased bioaerosol exposure, which is associated with poor lung function and more PEx among pwCF.

1.7 Aims and objectives

The aims of my Ph.D. were two-fold: first, I assessed the influence of ubiquitous environmental air pollutants such as NO_2 and $PM_{2.5}$ on disease severity in pwCF; and second, I wanted to better understand the influence of where pwCF live on disease severity and prevalence of bacterial and fungal infections for example by looking at proximity to composting facilities, a proxy for bioaerosol exposure in the UK. To do this, the three core chapters of my thesis had the following objectives:

- To estimate the effect of NO_2 and $PM_{2.5}$ on disease severity including lung function among children and adults with CF in Great Britain (GB). (Chapter II)

- To estimate the effect of $PM_{2.5}$ exposure on longitudinal changes in lung function among a cohort of pwCF with a residential postcode in Greater London over 10 years (2008-2017). (Chapter III)
- To examine the potential associations between distance to permitted composting site (PCS), a bioaerosol exposure proxy, and changes in lung function, PEx, bacterial and fungal infections in pwCF. (Chapter IV)

The choice of study design for each objective was meticulously considered, balancing data availability with the research objectives. While longitudinal designs offer powerful advantages for analysing temporal trends and causal relationships, their feasibility hinges on data accessibility and project timelines. Cross-sectional studies provide invaluable insights where data limitations constrain the scope of analysis.

Study I: Short-term (one year) exposure to air pollutants and disease severity in Great Britain: High-resolution annual air pollution estimates were available for Great Britain for 2015, while the best clinical data at the time of analysis were available for 2016. As the temporal windows of these datasets overlapped, a cross-sectional design provided the most reliable opportunity to assess the association between air pollution exposure and disease severity in pwCF.

Study II: Long-term (ten years) exposure to air pollutants and disease severity in the Greater London Area: Prioritizing the impact of long-term air pollution exposure on disease progression, I used a longitudinal design. Fortunately, air pollution data spanning ten years (2008-2017) was accessible for the Greater London Area. This temporal coverage, coupled with clinical data on London-based pwCF, made a longitudinal design both feasible and optimal for exploring long-term associations between air pollutants and disease severity in pwCF living in the Greater London Area.

Study III: Bioaerosol exposure and disease severity in the UK: Because the best clinical data available at the time of analysis were for 2016 and national data on composting sites were available for the whole of the UK, a cross-sectional design was employed across the UK. This approach permitted assessment of the national correlation between the impact of composting facilities on respiratory health and disease severity among pwCF in the UK.

2 Chapter II: Long-term (ten years) exposure to air pollutants and disease severity in the Greater London Area

2.1 Background

NO_2 and $PM_{2.5}$ are air pollutants associated with a range of health effects, including cancer, stroke, dementia, diabetes, respiratory and cardiovascular diseases (EEA, 2021; Holgate *et al.*, 2016). According to WHO estimates ambient air pollution contributes to deaths in 29% of lung cancer, 24% of stroke, 25% of heart disease, and 43% of lung disease deaths worldwide (WHO., 2024). These two pollutants are the most prevalent, widespread, and extensively studied air pollutants, and are associated with poor health outcomes in the general population (Schraufnagel *et al.*, 2019a; Schraufnagel *et al.*, 2019b). They have been recognised as major pollutants because they can infiltrate the lungs and produce an inflammatory reaction (Dauchet *et al.*, 2018; WHO., 2016). According to the European Environment Agency (EEA), nearly 300,000 and 40,000 premature deaths in the 27 EU Member States were attributed to $PM_{2.5}$ and NO_2, respectively, in 2019 (EEA, 2021).

Poor health outcomes linked to air pollution are more prevalent in vulnerable population sub-groups, including the elderly, people from low-income groups, and individuals with respiratory diseases such as COPD (Yang *et al.*, 2021a), asthma (Pfeffer *et al.*, 2021) and CF (Nieuwenhuijsen *et al.*, 2018; Wang *et al.*, 2017). Higher levels of NO_2 and $PM_{2.5}$ have a stronger impact on these vulnerable sub-populations, increasing the risk of cytotoxicity to inflammatory processes and thereby increasing morbidity and mortality (Gurram *et al.*, 2019; Hajat *et al.*, 2015; Taylor-Robinson & Schechter, 2011).

2.1.1 Impact of air pollution (exposures) on CF

The bronchial epithelial cells of people with pwCF are highly sensitive to oxidative stress caused by air pollution. Exposure to NO_2 and $PM_{2.5}$ results in oxidative stress and apoptosis in pwCF's bronchial epithelial cells at a much lower level of air pollution than in healthy bronchial epithelial cells (Kamdar *et al.*, 2008). Air pollutants trigger the cellular response in the lung and result in intense oxidative and nitrosative stress that increases the susceptibility to inflammation and infection in pwCF (Jassal *et al.*, 2013). The CF lung has low levels of S-nitrosoglutathione and high levels of free myeloperoxidase and neutrophil-derived proteases

(Berger, 2002; Snyder *et al.*, 2002). Exposure to air pollution may cause further oxidative stress, inflammation, and damage to the airways at low levels of S-nitrosoglutathione and high levels of myeloperoxidase. This could potentially affect the severity of the disease in pwCF.

CF bronchial epithelial cells produce less IL-6, IL-8, and MCP-1 in response to exposure to gasoline exhaust, whereas non-CF cells produce more of these pro-inflammatory mediators (Künzi *et al.*, 2015). NO_2 affects the CF lungs in two ways: first, by damaging the lung tissue through direct exposure; and second, indirectly by producing a systemic inflammatory response, increasing oxidative stress and decreasing the ability of microbial clearance (Schraufnagel *et al.*, 2019a). Exposure to NO_2 causes the release of proinflammatory cytokines in airway epithelial cells, which results in airway diseases (Blomberg *et al.*, 1999). The underlying mechanism by which NO_2 causes diseases is not known, but inflammation and oxidative stress associated with NO_2 may play an important role (Persinger *et al.*, 2002; Sandstrom *et al.*, 1991). NO_2 irritates the airway of the lungs and short-term exposure to NO_2 could lead to coughing, wheezing and breathing difficulties, while long-term exposure may contribute to the development of chronic respiratory disease and increase the susceptibility to infection in the lungs, which may contribute to premature death (Huangfu & Atkinson, 2020).

A study by Kamdar *et al.* examined the impact of $PM_{2.5}$ on the epithelium of pwCF by cultivating IB3-1 and S-9 CF human bronchial epithelial cells (Kamdar *et al.*, 2008). They discovered that $PM_{2.5}$ causes oxidative stress by mediating apoptosis in the bronchial epithelium of pwCF and that the high inflammatory response and impaired mitochondrial function caused by the pollutant play a significant role in the cytotoxic effects on the airway epithelium of pwCF. In addition to a study comparing the deposition and retention of technetium-99m-labelled iron oxide particles (5 microns) in pwCF and healthy volunteers, Brown *et al.* found a higher coarse particulate deposition in the tracheobronchial regions of poorly ventilated lungs in pwCF ($p<0.001$) (Brown *et al.*, 2001). $PM_{2.5}$ has been demonstrated to increase mitochondrial-signalled apoptosis in CF bronchial epithelial cells via caspase-9 and Bcl family proteins *in vitro* at significantly lower levels than in healthy controls, which may compromise pathogen clearance (Kamdar *et al.*, 2008; Psoter *et al.*, 2016).

2.1.2 Health outcomes in CF

To get a comprehensive overview of the current evidence on health outcomes and the main air pollutants in CF, I searched two databases: Ovid Medline and Embase and Web of Science. The time period ranged from January 1st, 1990, to December 31st, 2022. The search terms used were: ‘‘cystic fibrosis” or“CF” or “*CFTR*” combined with the specific terms ‘‘outdoor air pollution”, or ‘‘NO_2”, or ‘‘NOx” or ‘‘elemental carbon”, or ‘‘black carbon”, or ‘‘nitrogen dioxide”, or ‘‘nitrogen oxide”, or ‘‘$PM_{2.5}$”, or ”Fine Particulate Matter”, or ‘‘particulate matter”, and "lung function", or "FEV_1", or “percent predicted FEV_1”, or “pulmonary exacerbation”. Keyword search was limited to title, abstract and humans. Of the 47 research studies retrieved from the above databases, only four looked at how NO_2 or $PM_{2.5}$ affected pwCF's lung function or PEx. Of those four research studies conducted in pwCF, two were from the US (Goss *et al.*, 2004; Jassal *et al.*, 2013), one was from Belgium (Goeminne *et al.*, 2013) and one was conducted in Brazil (Farhat *et al.*, 2013). This further highlighted the current lack of evidence on this topic from the UK.

2.1.3 Air pollution and CF lung function (ppFEV$_1$)

Lung function is a key indicator of respiratory health as well as a predictor of cardiovascular and respiratory morbidity and mortality. Although about 50% of the variation in lung function among pwCF has been attributed to environmental factors, little is known about the influence of NO_2 and $PM_{2.5}$ on the health of pwCF (Collaco *et al.*, 2010). I found only one study that explored the impact of NO_2 and $PM_{2.5}$ on the lung function of pwCF (Goss *et al.*, 2004). In that study of 11,484 pwCF in the United States, Goss *et al.* found a 24 ml (95% CI; 7 - 40) drop in FEV_1 for every 10 ug/m^3 increase in $PM_{2.5}$ (Goss *et al.*, 2004). Similarly, there was a 0.5% (95% CI, 0.3–0.9%) reduction in the mean ppFEV$_1$ for each 10 $\mu g/m^3$ increase in $PM_{2.5}$ in the same study. However, Goss *et al.* found no association between NO_2 and lung function in pwCF.

2.1.4 Air pollution and CF pulmonary exacerbations (PEx)

PEx are harmful to pwCF, as they are associated with disease progression, poor quality of life, and an increased risk of mortality (de Boer *et al.*, 2011; Sanders *et al.*, 2011). In pwCF, PEx and poor lung function were attributed to a high concentration of air pollutants (Blayac *et al.*,

2022). It has also been reported that the odds of requiring an additional antibiotic course among pwCF are higher with an increase in air pollution (Goeminne *et al.*, 2013).

2.1.4.1 $PM_{2.5}$ and PEx

The relationship between $PM_{2.5}$ and PEx in the CF population was investigated in two studies. The first study conducted by Goss *et al.* in 2004 reported a 21% (95% CI: 7%;33%) increase in the odds of two or more exacerbations for every 10 µg/m^3 increase in $PM_{2.5}$ (Goss *et al.*, 2004). This estimate was adjusted for age, sex, race, weight, airway colonisation, pancreatic function, and insurance status in the regression model.

The other study was conducted by Jassal *et al.* on 145 pwCF from the CF centre at Children's Hospital Los Angeles (CHLA) in the US from 2004 to 2008 using annual air pollutant levels that were obtained from US Environmental Protection Agency monitoring stations (Jassal *et al.*, 2013). Individual exposure to $PM_{2.5}$ was assigned based on residential proximity to the nearest monitoring station. They found no association between CF exacerbations and $PM_{2.5}$ levels of 12.0 µg/m^3 (OR = 1.25, 95% CI: 0.59; 2.67, p= 0.555). Age, sex, ethnicity, insurance, genotype, pancreatic insufficiency, microbial flora, and pulmonary function decline were all factored into the odds ratios.

2.1.4.2 NO_2 and PEx

Three investigations examined the effect of NO_2 on PEx in pwCF: Goss *et al.* in the US, Goeminne *et al.* in Belgium, and Farhat *et al.* in Brazil. Goss *et al.* conducted the first and largest study, which discovered that for every 10 µg/m^3 increase in $PM_{2.5}$, there was a 1.21% (95%CI: 1.07; 1.33%) increase in the odds of two or more exacerbations. Details of the study were described in the previous section: air pollution and CF lung function (Goss *et al.*, 2004). The second study by Goeminne *et al.* used a case cross-over study design to examine the relationship between PEx and air pollution among 215 pwCF (both children and adults) between 1998 and 2010 in Belgium (Goeminne *et al.*, 2013). Goeminne *et al.*used a 4x4 km grid to allocate exposure to air pollutants. They discovered that the median interquartile range (IQR) NO_2 level was 24 (17-32) µg/m^3, and the probability of starting antibiotics increased by 1.106 (95%CI: 1.05; 1.17) for every 10 µg/m^3 rise in NO_2. The third study by Farhat *et al.* reported a significantly increased risk of PEx (relative risk = 1.86; 95%CI: 1.14;

3.02) for a 45.62 μg/m^3 IQR increase of O_3 concentration in a study of 103 pwCF with a median age of 8.9 years in the city of São Paulo, Brazil, between 2006 and 2007 (Farhat *et al.*, 2013).

2.1.5 The rationale of the study

Air pollution has a significant impact on the onset and progression of respiratory diseases. However, less is known about the influence of common air pollutants on the health of pwCF living in the UK, compared to other respiratory diseases. The wider role of environmental factors in the severity of CF has been largely neglected, with very limited data from the UK. Understanding how NO_2 and $PM_{2.5}$ contribute to disease progression and severity in the CF population will help in the development of more efficient ways and technologies for treating NO_2 and $PM_{2.5}$-induced diseases, as well as in the prevention and earlier diagnosis of the associated health problems in the UK. Improving air quality has substantial, measurable and important public health benefits and knowledge of the impact of air pollutants on lung function and PEx could help to design targeted and precise environmental health interventions that will improve the overall respiratory health of pwCF in the UK.

2.1.6 Hypothesis

I hypothesise that high mean annual NO_2 and $PM_{2.5}$ concentrations are associated with a decrease in lung function and an increase in PEx among pwCF.

2.1.7 Objective of the study

I aim to estimate the effect of NO_2 and $PM_{2.5}$ on disease severity including lung function defined by $ppFEV_1$ and PEx defined by the number of IV days (#IVdays) on antibiotics for PEx among children and adults with CF in Great Britain (GB).

2.2 Materials and methods

2.2.1 Study design

Using retrospective cohort data from the UK CF Registry, I used a semi-individual cross-sectional study design to quantify the impact of two major air pollutants on $ppFEV_1$ and PEx in pwCF in GB.

2.2.2 Study geography

The analysis was conducted at the pwCF's residential full postcode level (postcode unit). In the UK, there are 1.7 million postcodes and approximately 30 million individual postal addresses. A full postcode is known as a "postcode unit" and it designates an area with an average of 43 (SD: 38, median=33) people and 18 (SD: 15, median=14) households (ONS., 2013). The UK CF Registry checks postcodes on a regular annual basis and updates its database. Geographic coordinates from the Code-Point Open dataset (Ordnance Survey., 2022) and pwCF's residential full postcodes from the 2016 annual report were linked using ArcGIS 10.7.0 (ESRI., 2018).

Using the open dataset (Code-Point Open) of geographic coordinates of all the current postcode units in Great Britain available at https://www.ordnancesurvey.co.uk/business-and-government/products/code-point-open.html (Ordnance Survey., 2022), I used Voronoi polygons to create a shapefile of postcode boundary polygons in ArcGIS (ESRI., 2018). These polygons are a systematic way of dividing a surface into a number of regions using point data (Yamada, 2016). For each postcode unit, there is a corresponding region called a Voronoi cell or polygon. Each polygon contains only a single postcode unit input feature, as shown in ***Figure 2.1*** below. These Voronoi polygons are also called Thiessen Polygons.

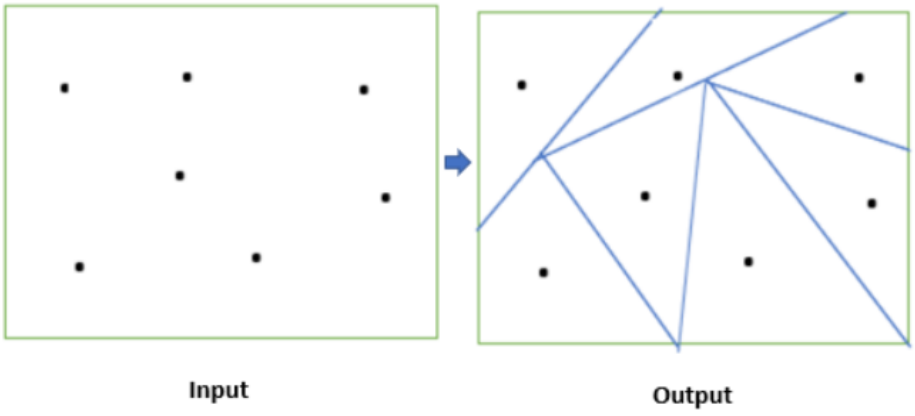


Figure 2.1: Conversion of individual postcode to Thiessen polygon boundary

These polygons, later termed "postcode polygons", represent approximate postcode unit boundaries in Great Britain. As the exact location of each CF individual within a postcode area is unknown; postcode area level polygons were created to improve local exposure measurement and account for uncertainties in exposure assignment. I conducted two

different analyses, in the first set of analyses, I used the geographic centroid of the postcode, and then I used the averaged exposure across the postcode area.

2.2.3 Study population

The UK CF Registry holds data on over 99% of pwCF in the UK (Taylor-Robinson *et al.*, 2018). There are over 10,000 patients from all CF specialist centres across the UK. These pwCF have an annual assessment from which data is entered into the UK CF Registry. Data are collected on the demography, health outcomes and social factors of individuals, with a report produced annually presenting key summaries of the collected data which can be accessed at (https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources). In the UK, pwCF transitions from paediatric to adult care around the age of 16 years. Individuals younger than 16 (0- 15) years are therefore categorised as children, and those aged 16 and above are adults.

This analysis includes all pwCF, who had an annual review in 2016, lived in Great Britain (GB) and registered in the UK CF Registry with data on age, sex, lung function, height, weight, BMI, number of exacerbations, *P. aeruginosa* and postcode. I used clinical data for the year 2016 because it was the most recent data available at the time this study was set up.

2.2.4 Exclusion criteria

PwCF who had a lung transplant in 2016 or before 2017 (n=293) were excluded as their transplanted lungs and immunocompromised state would confound results.

2.2.5 Outcome variable

I focused on two key measures commonly used by clinicians to assess disease severity in CF: i) ppFEV_1 as a measure of lung function, and ii) the number of PEx per year. I used these two measures independently and in combination throughout this analysis as proxies of clinical severity in CF.

2.2.6 Lung function

Lung function was measured as percent predicted (pp)FEV_1, as this is a key prognostic measure in CF (Szczesniak *et al.*, 2017a). FEV_1 is the amount of air an individual can blow out in one second and is also an important measure of health status and treatment response in

CF (Szczesniak *et al.*, 2017a) $ppFEV_1$ is estimated using the Global Lung Function Initiative prediction equation, which takes into account factors including age, sex, height, and ethnicity in FEV_1 (Quanjer *et al.*, 2012). I used $ppFEV_1$ as a continuous variable.

Spirometry, the method used to assess lung function, is not a reliable measure for children under the age of six years due to the child's inability to perform standard pulmonary function tests accurately. Therefore, all pwCF younger than six years old (n= 1,494) were excluded from our $ppFEV_1$ analysis.

2.2.7 Pulmonary exacerbation

There is no standard definition of a PEx. However, the use of intravenous (IV) antibiotics (which are used to treat a PEx) is widely used as a proxy marker in observational studies. Here, a PEx was defined as the use of IV antibiotics at home or in the hospital. I used the total number of days on IV antibiotics (home and hospital IV courses) received for a PEx during the twelve months preceding the annual review as I considered it to be a more robust and reliable measure of disease severity than the number of antibiotic courses as it is recorded accurately in the UK CF registry (Hoo *et al.*, 2018). Essentially, even though the 'standard' course length is 14 days, people may experience substantially longer (or occasionally shorter) courses; this is mostly dependent on the severity of the disease at baseline and the severity of the exacerbation. For instance, if the analysis is limited to the number of PEx courses (i.e., one PEx episode for both), a patient with a 21-day course will have the same data as a patient with a 14-day course. However, it is highly likely that the patient with the longer course is more ill (and therefore more severe), so the number of days is more likely to be a better indicator of the severity of the condition. Additionally, although oral antibiotics may also be used to treat a PEx, the data are less reliable as courses are sometimes given outside of a CF centre and the patient may not always report these.

2.2.8 Exposure variables and covariates

To capture spatial differences in the effects of environmental factors such as air quality, it is important to link the clinical characteristics of each pwCF to NO_2 and $PM_{2.5}$ model estimates, while adjusting for various covariates. I linked the pwCF with NO_2 and $PM_{2.5}$ maps using the residential postcode of each pwCF.

Environmental factors (e.g., air pollution exposure) were linked to the clinical characteristics of each pwCF using their residential postcode. For every pwCF, their postcode was saved on the Small Area Health Statistics Unit (SAHSU) secure server in a separate file from their clinical data with restricted and controlled access in compliance with UK CF Registry standards, and data confidentiality and privacy laws. These postcodes were the key linking variable; therefore, I verified they were error-free before linking them with air pollution maps. The CF Registry management team also confirmed that these postcodes are verified annually during the pwCF's annual review visit to the CF centre. There was a separate linkage file for individual postcodes and clinical data, which is in accordance with the following recommended practice (i.e. to keep a separate linkage and the analysis files to maintain confidentiality (Harron *et al.*, 2017).

2.2.9 Air quality

Many studies assign air pollution exposure to individuals from the nearest monitors, but often the monitoring networks are not dense enough to capture the spatial heterogeneity in ambient air pollution among individuals living in large geographic areas. Therefore, in my study, NO_2 and $PM_{2.5}$ exposures were extracted from air pollution surfaces developed from detailed air pollution models. I used two different high-resolution model outputs to assign exposures to each pwCF. I started with a high-resolution (100m X 100m) continuous modelled surface for NO_2 for the year 2009 (Gulliver *et al.*, 2013) and $PM_{2.5}$ for the year 2010 (de Hoogh *et al.*, 2016) in Great Britain available from SAHSU. The NO_2 and $PM_{2.5}$ surfaces developed by Gulliver *et al.* and de Hoogh *et al.* were based on land-use regression (LUR) (de Hoogh *et al.*, 2016; Gulliver *et al.*, 2013). LUR modelling is a statistical approach widely used for estimating spatial variation in air pollution and measuring population air pollution exposure. In environmental epidemiological studies, LUR modelling is increasingly being utilised to predict spatial and temporal variation in ambient air pollution exposures for current or recent periods (Boniardi *et al.*, 2019; Clark *et al.*, 2010; Eeftens *et al.*, 2016; Ma *et al.*, 2020). LUR models provide a relatively robust method for spatial prediction while using less data than dispersion models and requiring less sampling effort than geostatistical models (Hoek *et al.*, 2008). These models also produce a better and higher spatiotemporal resolution map for estimating historical air pollutant exposure and are also useful for estimating air pollution concentrations where input data are limited and scarce.

LUR models are based on multivariate regression equations using the monitored levels of the pollutant(s) of interest as the dependent variable and relevant environmental variables including traffic, population density, land use, physical geography (e.g. altitude), wind speed, climate and satellite imagery usually obtained through geographic information systems (GIS) as the independent variables (de Hoogh *et al.*, 2016; Gulliver *et al.*, 2013). The regression equations are then used to predict the concentration at an unmeasured location based on the model. More details about the LUR Model can be found in de Hoogh *et al.* (de Hoogh *et al.*, 2016). This LUR model explained almost 60% of the spatial variation in measured $PM_{2.5}$ concentrations and 58% in NO_2 levels. Similarly, cross-validated coefficient of determination (CV R^2) values for $PM_{2.5}$ and NO_2 were .0.59 and 0.55, respectively (de Hoogh *et al.*, 2016). CV R^2 was obtained by regressing estimated concentrations against measurements, the range of CV R^2 is between 0 and 1, and the closer it is to 1 the better it is, meaning the estimates are closer to the actual measurements.

A higher-resolution (25m X 25m) continuous modelled surface for annual average concentrations of NO_2 and $PM_{2.5}$ in Great Britain in 2015 was developed within SAHSU during my PhD (Wang *et al.*, 2022b). I managed to get access to these surfaces, which were more contemporary to the UK CF Registry data used in this analysis. The purpose of developing these models was to predict annual pollution concentrations on a national scale. These nationally modelled surfaces were developed using LUR techniques and spatiotemporal models in a generalised additive mixed model framework to estimate NO_2 and $PM_{2.5}$ at 25 m grid cells. The model was based on a broad range of predictors, including land cover, road network, traffic, topography, buildings, census data, and the number of households per postcode. The models were cross-validated against measurements at the monitor stations of the Automatic Urban and Rural Network (AURN). The models achieved CV R^2 of 0.75 and 0.62 for NO_2 and $PM_{2.5}$, respectively (Wang *et al.*, 2022b). ***Figure 2.2*** shows the modelled NO_2 and $PM_{2.5}$ concentrations in GB.

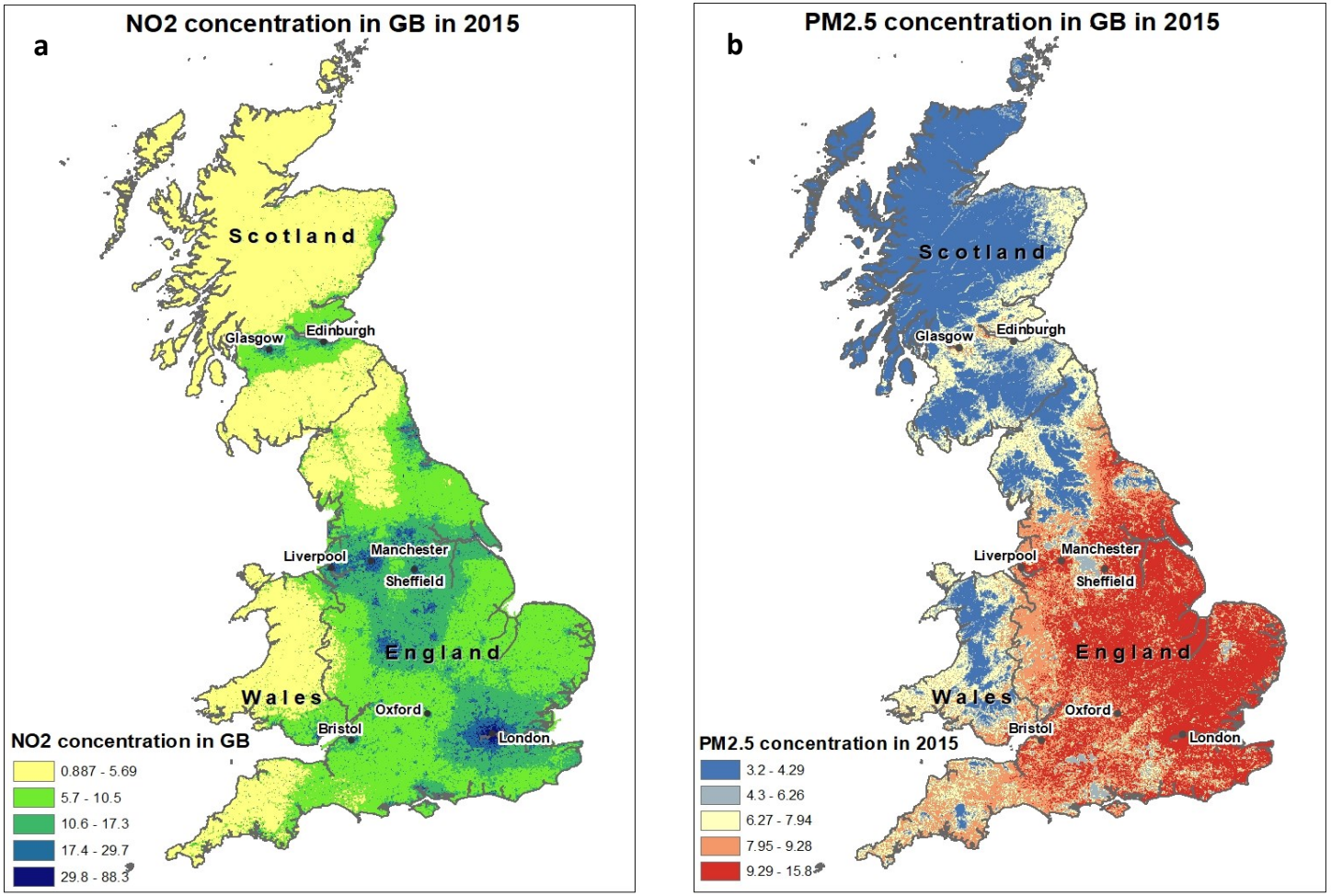


Figure 2.2: a: NO_2 concentrations in GB 2015, b: $PM_{2.5}$ concentrations in GB 2015

Using ArcGIS 10.7.1 (ESRI., 2018), I have extracted the NO_2 and $PM_{2.5}$ concentrations from both sets of models at each postcode (Analysis 1) and postcode polygon (Analysis 2) of all pwCF included. The ArcMap tool "Extract raster values to point locations" was used to determine the NO_2 and $PM_{2.5}$ exposure estimates at the residential postcode point of pwCF. The average exposure to NO_2 and $PM_{2.5}$ at the postcode area level was determined for every postcode polygon defined in the study geography in Section 2.2.2 using the ArcGIS Zonal Statistics tool. The former resulted in a single value, while the latter included the mean, range, minimum and maximum concentrations, allowing me to assess the importance of potential exposure misclassification.

2.2.10 Covariates

In pwCF, a number of factors like age, sex, BMI, *CFTR* genotype, rural-urban classification, and socio-economic position (deprivation), are known to influence lung function, IV antibiotic use, and overall disease severity (Er *et al.*, 2020; Kerem *et al.*, 2014; Taylor-Robinson *et al.*, 2013). It was, therefore, important to explore and adjust for their effect in this semi-individual analysis. Age, sex, body mass index (BMI), *CFTR* genotype, *P. aeruginosa* sputum culture, urban-rural classification, deprivation, and nation were the *a priori* variables that I identified as potential confounders. A justification for each of these covariates is provided below.

2.2.10.1Age

FEV_1 increases from childhood until the age of 16 (Talaminos Barroso *et al.*, 2018), then drops with age due to decreased chest wall compliance, decreased expiratory muscular strength, and a narrower airway during forced expiration (Lalley, 2013), before decreasing faster in those aged 40 and above. As a result, I originally divided the age variable into four categories: 0 to 5, 6 to 16, 16 to 40, and 40 and above. Due to the small number of people in the 40 and above group, I combined it with the 16 to 40 group.

2.2.10.2 Sex

When compared to women of the same height and weight, men have larger lungs with more bronchi, a larger alveolar surface area, and a broader airway (Townsend *et al.*, 2012). As a result, men have higher ventilation, peak inspiratory pressure, and FEV_1 than women. There

are also known gender disadvantages for females with CF (Verma *et al.*, 2005). In this investigation, I used a binary variable (M *vs* F) to adjust for these differences.

2.2.10.3 Body mass index (BMI)

BMI is calculated as weight divided by height squared and is one of the key measures used to monitor and evaluate changes in the nutrition status of an individual in clinics or hospitals. Zemel *et al.* reported a significant association between nutritional status and lung function among CF children (Zemel *et al.*, 2000). A BMI over the 50th percentile for children and an absolute BMI >19 kg/m^2 for adults is linked to better CF outcomes (Stenbit & Flume, 2011). As a result, I used absolute BMI in adults and BMI percentile, calculated based on age, sex, height and weight, in children using the R package "*childsds*" (Vogel M., 2022.). I accounted for BMI in both my ppFEV$_1$ and PEx analyses.

2.2.10.4 *CFTR* genotype

I developed, with input from CF clinicians, a new classification, with the aim of making it more relevant to contemporary practice. A new eight-group classification system was devised, using terminology developed for new *CFTR*-targeted therapies – so-called *CFTR*-modulators: residual function (RF), minimal function (MF) and gating function (Heijerman *et al.*, 2019; Middleton *et al.*, 2019) in combination with the *F508del* categorisation (Clinical and Functional Translation of CFTR (CFTR2). 2022a). In addition to better capturing the heterogeneity of known phenotypic variability in CF, I believe that one of the additional benefits of this new classification is that it aligns with prospective therapeutic options for patients, which is becoming increasingly important in clinical care and outcomes. Each of the eight categories is described below: This new classification of *CFTR* genotypes is summarised in ***Table 2.1*** and each group is briefly explained in more detail below.

Table 2.1: Classification of *CFTR* genotypes with frequencies among pwCF (n=9,361) in the UK in 2016

Group	Mutations	Description	n	%
1	**Homozygous *F508del***	*F508del* on both alleles.	4,603	49.2
2	***F508del*/ Minimal function (MF)**	One allele is *F508del* and the other one is any minimal function mutation i.e., a mutation that produces either no protein or protein that does not	1,822	19.5

		respond to currently approved *CFTR* modulators,		
3	**Residual function (RF)/other**	RF mutations include 2789+5G>A, 3272-26A>G, 3849+10kbC>T, 711+3A->G, A1067T, A455E, D110E, D110H, D1152H, D1270N, D579Gx, E193K, E56K, E831X, F1052V, F1074L, G1069R, K1060T, L206W, P67L, R1070Q, R1070W, R117C, R347H, R352Q, R74W, S945L and S977F. RF on one allele and any other mutation on the 2nd allele were counted in this group including RF/RF, *F508del*/RF, MF/RF, RF/Gating and RF/Other. This group also includes R117H.	1,070	11.4
4	Gating (G)/other	G551D, G1244E, G1349D, G178R, G551S, S1251N, S1255P, S549N, or S549R mutations are classified as gating mutations. Gating on one allele and any other mutation on the 2nd allele including *F508del*/G, G/G, G/Other and MF/G <u>but not</u> RF/G (see above). Other = not identified or identified but classification into a specific group was not possible.	583	6.2
5	**MF/MF**	All genotypes with MF mutations on both alleles	298	3.2
6	***F508del*/Other**	*F508del* on one allele and the second allele was either not identified or identified but classification into the above groups was not possible	663	7.1
7	**MF/other**	MF on one allele and the second allele was either not identified or identified but classification into the above groups was not possible	146	1.6
8	**Other/other**	On both alleles either the mutation was not identified, or it was identified but classification into the above categories was not possible.	176	1.9

2.2.10.4.1 Homozygous *F508del*

F508del represents 70% of all mutant *CFTR* genes worldwide and in the UK, 90% of pwCF have at least one *F508del* mutation, and 50% have two; it corresponds to a phenylalanine deletion at position 508 of the *CFTR* protein (Fiore *et al.*, 2020; Lopes-Pacheco, 2020). With this mutation, the mutant protein is produced but fails to fold properly inside the cell, resulting in a lack of

mature *CFTR* at the cellular membrane to perform its specific chloride ion conductance function (Lukacs & Verkman, 2012).

2.2.10.4.2 *F508del*/ Minimal function (MF)

MF mutations are those that cause the *CFTR* protein to be minimally functional or not functional at all (Lopes-Pacheco, 2020). Two sources were used to identify mutations associated with minimal function. First, using the list (n=220) of minimal function mutations available at (https://www.cff.org/PDF-Archive/Study-VX18-445-106-Eligible-Mutations-June-2019.pdf) (Clinical and Functional Translation of CFTR (CFTR2). 2022a). The mutation was classed as minimal if it matched the list included in ***Table 2.2***. As the position of the alleles does not affect the outcome, the order of the two mutations is not relevant. Permutations *(F508del/Q2X and Q2X/F508del*) were therefore considered in the same way.

Table 2.2: List of minimal function mutations

Q2X	*L218X*	*Q525X*	*R792X*	*E1104X*
S4X	*Q220X*	*G542X*	*E822X*	*W1145X*
W19X	*Y275X*	*G550X*	*W882X*	*R1158X*
G27X	*C276X*	*Q552X*	*W846X*	*R1162X*
Q39X	*Q290X*	*R553X*	*Y849X*	*S1196X*
W57X	*G330X*	*E585X*	*R851X*	*W1204X*
E60X	*W401X*	*G673X*	*Q890X*	*L1254X*
R75X	*Q414X*	*Q685X*	*S912X*	*S1255X*
L88X	*S434X*	*R709X*	*Y913X*	*W1282X*
E92X	*S466X*	*K710X*	*Q1042X*	*Q1313X*
Q98X	*S489X*	*Q715X*	*W1089X*	*Q1330X*
Y122X	*Q493X*	*L732X*	*Y1092X*	*E1371X*
E193X	*W496X*	*R764X*	*W1098X*	*Q1382X*
W216X	*C524X*	*R785X*	*R1102X*	*Q1411X*
185+1G>T	*711+5G>A*	*1717-8G>A*	*2622+1G>A*	*3121-1G>A*
296+1G>A	*712-1G>T*	*1717-1G>A*	*2790-1G>C*	*3500-2A>G*
296+1G>T	*1248+1G>A*	*1811+1G>C*	*3040G>C (G970R)*	*3600+2insT*
405+1G>A	*1249-1G>A*	*1811+1.6kbA>G*	*3850-1G>A*	*4005+1G>A*
405+3A>C	*1341+1G>A*	*1811+1643G>T*	*3120G>A*	*4374+1G>T*
406-1G>A	*1525-2A>G*	*1812-1G>A*	*3120+1G>A*	*3791delC*
621+1G>T	*1525-1G>A*	*1898+1G>A*	*3121-2A>G*	*3821delT*
711+1G>T	*1898+1G>C*	*1782delA*	*2732insA*	*3876delA*
182delT	*1119delA*	*1824delA*	*2869insG*	*3878delG*
306insA	*1138insG*	*1833delT*	*2896insAG*	*3905insT*
365-366insT	*1154insTC*	*2043delG*	*2942insT*	*4016insT*

394delTT	*1161delC*	*2143delT*	*2957delT*	*4021dupT*
442delA	*1213delT*	*2183AA>G a*	*3007delG*	*4022insT*
444delA	*1259insA*	*2184delA*	*3028delA*	*4040delA*
457TAT>G	*1288insTA*	*2184insA*	*3171delC*	*4279insA*
541delC	*1343delG*	*2307insA*	*3171insC*	*4326delTC*
574delA	*1471delA*	*2347delG*	*3271delGG*	
663delT	*1497delGG*	*2585delT*	*3349insT*	
849delG	*1548delG*	*2594delGT*	*3659delC*	
935delA	*1609del CA*	*2711delT*	*3737delA*	
1078delT	*1677delTA*	*991del5*		
CFTRdele1	*CFTRdele16-17b*	*1461ins4*		
CFTRdele2	*CFTRdele17a,17b*	*1924del7*		
CFTRdele2,3	*CFTRdele17a-18*	*2055del9>A*		
CFTRdele2-4	*CFTRdele19*	*2105-2117del13insAGAAA*		
CFTRdele3-10,14b-16	*CFTRdele19-21*	*2372del8*		
CFTRdele4-7	*CFTRdele21*	*2721del11*		
CFTRdele4-11	*CFTRdele22-24*	*2991del32*		
CFTR50kbdel	*CFTRdele22,23*	*3121-977_3499+248del2515*		
CFTRdup6b-10	*124del23bp*	*3667ins4*		
CFTRdele11	*306delTAGA*	*4010del4*		
CFTRdele13,14a	*602del14*	*4209TGTT>AA*		
CFTRdele14b-17b	*852del22*	*Y569D*		
A46D	*V520F*	*L1065P*		
G85E	*A559T*	*R1066C*		
R347P	*R560T*	*L1077P*		
L467P	*R560S*	*M1101K*		
I507del	*A561E*	*N1303K*		

Second, there were ≥ 900 MF mutations defined by Vertex (i.e., eligible for Kaftrio). The company has many approved medicines to treat the underlying cause of CF, as well as several ongoing clinical and research programmes in the field. I used data made available online by Vertex (https://www.cfsource.co.uk/hcp) to define MF mutations. On their web portal, there are two search boxes for mutation 1 and mutation 2. So, to determine if it is a minimal function mutation, I used *F508del* in one search box and the mutation in question in the other. A mutation was only considered MF if the outcome was 'Kaftrio' (not Symkevi). *F508del* on one allele and any MF mutation on the other allele (as defined by the above criteria) were classified as *F508del*/MF.

2.2.10.4.3 Residual function (RF)

RF mutations affect about 5% of the CF population and result in residual *CFTR* ion transport due to partial *CFTR* expression and variable preserved channel gating or function (Rowe *et al.*, 2017; Salvatore *et al.*, 2019). These mutations are associated with delayed onset and slower progression of lung disease, though the disease may become severe in some patients, as well as pancreatic sufficiency and sweat chloride concentrations sometimes below the diagnostic threshold of 60 mmol/L (Salvatore *et al.*, 2019; Van Goor *et al.*, 2014).

A mutation was defined as RF if it matched the following list (Salvatore *et al.*, 2019). We included the mutation *R117H+5T* in this group as it is widely accepted to be associated with RF, even though it is often reported separately. This is because of its relatively high allele frequency which stimulated an early interest in *CFTR* modulator trial development (Moss *et al.*, 2015). The list used for my work is included in ***Table 2.3***.

Table 2.3: List of RF mutations

E56K	*R352Q*	*2789+5G>A*	*F1052V*
P67L	*A455E*	*3272-26A>G*	*G1069R*
D110H	*711+3A->G*	*3849+10kbC>T*	*R1070Q*
R117C	*E831X*	*R74W*	*R1070W*
E193K	*S945L*	*D110E*	*F1074L*
R347H	*K1060T*	*D579G*	*D1152H*
L206W	*A1067T*	*S977F*	*D1270N*
R117H			

Any genotype including an RF mutation on one allele was classified as RF. This group, therefore, includes the following combinations: RF/RF, MF/RF, Gating/RF and other/RF (including *F508del*/RF).

2.2.10.4.4 Gating

CFTR gating mutations are caused by mutations that result in *CFTR* protein which is present at the cell surface but does not open and close properly (defective channel gating), resulting in reduced chloride transport as measured *in vitro* or *in vivo* (Bobadilla *et al.*, 2002; Illek *et al.*, 1999). These missense *CFTR* mutations are found in between 4% and 8% of pwCF. *G551D, G1244E, G1349D, G178R, G551S, S1251N, S1255P, S549N*, or *S549R* mutations are classed as

Gating mutations (De Boeck *et al.*, 2014; Yu *et al.*, 2012). The most common *CFTR* gating mutation in pwCF is *G551D* (Yu *et al.*, 2012).

Gating on one allele and any other mutation on the 2nd allele were counted in this group including *F508del*/Gating, Gating/Gating, and MF/Gating but excluding RF/Gating.

2.2.10.4.5 MF/MF

All mutations associated with MF (as defined above) on both alleles were classified as MF/MF.

2.2.10.4.6 F508del/Other (Unclassified/not identified)

If there was *F508del* on one allele and the mutation on the second allele was either unclassified or unidentified, pwCF were classified as *F508del*/other.

2.2.10.4.7 MF/other

If there was MF on one allele and the mutation on the second allele was either unclassified or unidentified, the individuals were classified as MF/other.

2.2.10.4.8 Other/other

All unclassified and unidentified mutations were placed in the Other/Other categories (including unclassified/unclassified and not identified/not identified).

A mutation that was discovered but with limited information available was labelled as unclassified or other. For example, c.*1021 1022DUPTCINEXON7*, *EXON4-10deleTION*, and *V392G* were unclassified mutations. Similarly, if no mutation was detected despite genotyping, it was classified as not identified.

2.2.10.5 Sensitivity analysis of new mutation classification

To both assess the impact of my new classification and facilitate comparisons with previous studies, I performed my analysis twice: one using the novel mutation classification created in this research project and a second using the simplified version of the classification based on *F508del* mutation status as the main discriminator – that is, homozygous *F508del*, heterozygous *F508del*, and 'other.'

2.2.10.6 Pseudomonas aeruginosa

P. aeruginosa is the most common respiratory pathogen in the CF population and is associated with higher morbidity and mortality (Bhagirath *et al.*, 2016; Govan *et al.*, 2007; Ryall *et al.*, 2008). Patients who are chronically infected with *P. aeruginosa* have significantly poorer lung function and a higher risk of death compared to non-infected *P. aeruginosa* patients (Schaedel *et al.*, 2002). I used records of positive *P. aeruginosa* cultures (sputum, cough/throat swab, or bronchoalveolar lavage) in the year preceding the 2016 annual review to define the presence of *P. aeruginosa*.

2.2.10.7 Rural and urban

There are significant variations in health status among individuals living in rural and urban areas. Overall better health and lower mortality are generally documented among rural residents, but conversely, there is also some evidence of poor health and high mortality in rural areas (Kyte & Wells, 2010). The association between urban areas and disease is complex and context dependent. Urban areas can both pose health risks and provide health benefits, such as improved access to health care. In CF, Grehn *et al.* found a significant association between *A. fumigatus* colonisation and urban life in the German CF population in 2018 (p= 0.004) (Grehn *et al.*, 2020).

Based on pwCF's residential postcodes, I extracted rural and urban area classifications from ONS postcodes (ONS., 2019). All urban major and minor conurbations (population of 10,000 or more) and urban cities are classified as urban areas in England and Wales, while rural towns, outskirts, villages, hamlets, and isolated residences are classified as rural areas. In Scotland, a rural area was defined as one with a population of fewer than 3,000 people, while an urban area was defined as one with a population of 3,000 and above.

2.2.10.8 Deprivation

Though CF is a genetic disease, its severity varies among various socioeconomic classes. This has been reported in the UK and the US, where individuals from the most disadvantaged socioeconomic class die at a younger age compared to individuals from the least disadvantaged socioeconomic class (Barr *et al.*, 2011; Schechter *et al.*, 2001). Similarly, another study in the UK found that CF children from the most deprived area of the Index of

Multiple Deprivations (IMD) had poorer lung function compared to children from the least deprived area (Taylor-Robinson *et al.*, 2013).

Deprivation is a multi-dimensional problem for which several indices have been developed. Here, I considered two different indices of deprivation frequently used in epidemiological studies in the UK: i) the UK Townsend index scores; and ii) the English IMD.

Townsend index scores

As a proxy for socio-economic status, I used the 2011 UK Townsend index deprivation scores at the output area level (highest resolution available) (ONS., 2020). The Townsend index scores are based on four census variables: unemployment, overcrowding, lack of a car, and lack of a home (ONS., 2020). These scores were then divided into quintiles to help see the geographic distribution. The first quintile (Q1) relates to the least deprived areas of the UK, whereas the fifth quintile (Q5) refers to the most deprived. Townsend index can be argued to be flawed in including non-car ownership. As many people, choose not to buy a car, regardless of economic background, although often due to inadequate availability of affordable parking especially in most parts of London. In contrast, access to a car is more or less essential in rural areas and areas with limited public transport. In areas where means are limited, residents have to sacrifice other costs to be able to buy and drive a car. Another issue with Townsend index's calculations is that they are based on census data that can be up to twelve years old and may be inaccurate at the time of the current study (Adams *et al.*, 2005). Data from the 2021 Census were released too recently to be included in this work. Nonetheless, the Townsend index deprivation scores are not an absolute measure of poverty or deprivation, but rather an indicator of how poor an area is in comparison to others.

English IMD scores

The English IMD is a more recent proxy of the socio-economic status across England. The English IMD score uses 37 indicators that include seven distinct domains: income, employment, health deprivation and disability, education skills and training, barriers to housing services, crime and the living environment. Some of the domains have more effects on deprivation; therefore, the weighting of these domains was accounted for based on their contribution to the calculation of the IMD score 2015. For example, income and employment were given more weight at 22.5% followed by health deprivation and disability and education

were given a weight at 13.5%. The weight given to barriers to housing and services and crime was 9.3%.

The English IMD score was calculated for every Lower Layer Super Output Area (LSOA) across England; there are 32,844 geographic LSOAs. A rank of 1 is given to the most deprived LSOA and a rank of 32,844 is given to the least deprived LSOA. Quintiles of IMD were used for this analysis and linked to each pwCF's residential postcode. Q1 referred to the least deprived whereas Q5 referred to the most deprived area of the quintile.

IMDs are calculated independently using different indicators in Scotland and Wales, making UK- or GB-wide studies more difficult. English IMD 2015 data were downloaded from http://opendatacommunities.org/data/societal-wellbeing/imd/indices and the ONS postcode directory (https://geoportal.statistics.gov.uk/items/ons-postcode-directory-february-2019) (ONS., 2019).

2.2.10.9 Country

The best standard of care for pwCF is often provided by CF Specialist treatment centres. However, there are significant regional differences in access to specialised CF services (Barr *et al.*, 2011; Walters *et al.*, 1994)**.** According to research by Mahadeva *et al.,* patients who were cared for at a large specialist CF centre had better lung function than those who received local care at non-specialist clinics (Mahadeva *et al.*, 1998). I sought to determine whether the $ppFEV_1$ and #IVdays of pwCF differed depending on whether they resided in England, Wales, or Scotland.

2.2.11 Spatial and statistical analysis

Spatial and statistical analysis was performed using ArcGIS 10.7.1 (ESRI., 2018) and *R* 4.2.1, respectively (R Core Team., 2022).

2.2.11.1 Descriptive analysis

A descriptive analysis of variables (demographic and clinical characteristics) was performed to explore the distribution of the data. Mean and standard deviation were calculated for continuous variables ($ppFEV_1$, #IVdays for PEx, BMI, NO_2 and $PM_{2.5}$) and frequencies for

categorical variables (age, sex, genotype, *P. aeruginosa,* deprivation (Townsend index and IMD), rural/urban and country).

Descriptive and correlation analysis were used to study the distribution of the dataset using mean, standard deviation, range and Pearson's correlation statistical test respectively. When an independent variable in a multiple regression analysis is highly correlated with one or more of the other covariates, multicollinearity exists. Multicollinearity may result in higher standard errors and less reliable coefficients (Allen MP., 1997). Pearson correlation coefficients were calculated for the $ppFEV_1$, #IVdays, NO_2 and $PM_{2.5.}$

Scatter plots were generated to explore the relationship between NO_2, $PM_{2.5}$, $ppFEV_1$ and #IVdays. Scatter plots are very useful as they not only show the values of individual data points but also describe the patterns of the data as a whole. Regression models for children and adults were created separately. Univariate analyses were conducted to explore associations between air pollutants and clinical outcomes among pwCF with residential postcodes in GB.

I considered air pollution exposure as a continuous variable however, upon preliminary analyses, the results were found to lack clarity and robustness. Studies have demonstrated that the association between air pollution and health effects may not always follow a linear pattern, and there may be threshold effects or nonlinear relationships (Nasari *et al.*, 2016; Samoli *et al.*, 2005; Yuan *et al.*, 2023). By categorizing exposure levels into quintiles, one can better capture potential threshold effects and identify critical exposure levels where health risks may significantly increase (Rehan & Munir, 2022). The decision to use quintiles for the exposure variable, as opposed to a continuous approach, was informed by the need for a pragmatic and clinically meaningful representation of air pollution exposure. Quintiles provide a discrete categorization of exposure levels and account for the distribution of the exposure variable, which may not be normal, facilitating a more interpretable and actionable framework for both clinicians and policymakers. Comparing risk across quantiles provides a clearer picture of how risk changes with increasing exposure levels. This stratification allows for the identification of potential threshold effects or nonlinear associations that might be obscured in a continuous model.

2.2.11.2 ppFEV$_1$ analysis

Multiple linear regression models were used to assess the independent association of NO_2 and $PM_{2.5}$ with ppFEV$_1$ while controlling for age, sex, BMI, genotype, *P. aeruginosa*, and deprivation (Townsend index). The models were adjusted in a forward-stepwise approach with covariates added one-by-one using the Akaike information criterion (AIC) (Burnham & Anderson, 2003) for model selection. The best model was chosen based on the lowest AIC. The AIC is a statistical tool for determining how well a model fits the data from which it was developed, and the best-fit model is the one that explains the most variation in outcome variables with the fewest number of independent variables. *p*-values < 0.05 were interpreted as statistically significant for all results. The estimated change in ppFEV$_1$ with 95%CI was reported for the quintile of NO_2 and $PM_{2.5}$ exposure.

2.2.11.3 PEx analysis

For my PEx analysis, I chose a zero-inflated negative binomial regression as there was a large over-dispersion and excess zero in #IVdays. Overdispersion was checked by comparing the variance (814.3 days) and mean (16.7 days) of the response variable. The excess zero was examined by comparing the observed and predicted zeros using the performance and *glmmTMB* functions in *R*.

Zero-inflated negative binomial (ZINB) regression was used to explore the independent association between NO_2, $PM_{2.5}$ and #IVdays for PEx adjusting for age, sex, BMI, ppFEV$_1$, genotype, *P. aeruginosa*, and deprivation for both children and adults. Incidence rate ratio (IRR) with 95%CI for #IVdays was reported for the quintile of NO_2 and $PM_{2.5}$ exposure.

2.2.11.4 Sensitivity analysis

I performed various sensitivity analyses to validate my findings. First, I performed sensitivity analyses on the ppFEV$_1$ and #IVdays data for the subset of study participants who didn't change their residence's postcode over a period of five years (2012-2016). Second, the mean FEV$_1$ by NO_2 and $PM_{2.5}$ in groups with severe and mild mutations were compared. The severe mutation group included homozygous *F508del*, *F508del*/MF, Gating/any (Except RF mutation), and MF/MF categories, whereas the mild mutation group included Residual (RF)/other, F508del/other, Other/Other, MF/other. Third, I substituted an IMD score for

Townsend index deprivation for a subgroup of English residents in exploring the effect of NO_2 and $PM_{2.5}$ on $ppFEV_1$. Finally, I performed subgroup analyses in patients, excluding outliers with over 180 and over 120 IV days, respectively, for the IV data analysis**.**

2.2.12 Ethical considerations

This study employed individual-level data that was pseudonymized by the UK CF Registry, which has permission from every individual whose data was gathered, with approval from the UK National Health Service (NHS) research ethics committee (REC ref: 07/Q0104/2). The UK CF Registry Research Committee gave its approval for the use of the data (Data Request Reference: 302). Data were held and processed within a safe environment which undergoes regular independent third-party audits for ISO/IEC 27001 compliance.

2.3 Results

2.3.1 Overview of the CF data

There were 9,695 pwCF recorded in the CF Registry who completed an annual review in 2016. Of these, 730 were excluded from our analyses for the following reasons: 14 had a missing postcode; 423 had a residential postcode outside of Great Britain and 293 had a lung transplant (***Figure 2.3***). As a result, 8,965 people were eligible for our PEx analysis. For our $ppFEV_1$ analysis, I excluded children under the age of six (n=1,494) for reasons explained in ***Section 2.2.6***. The number of pwCF considered was therefore 7,471.

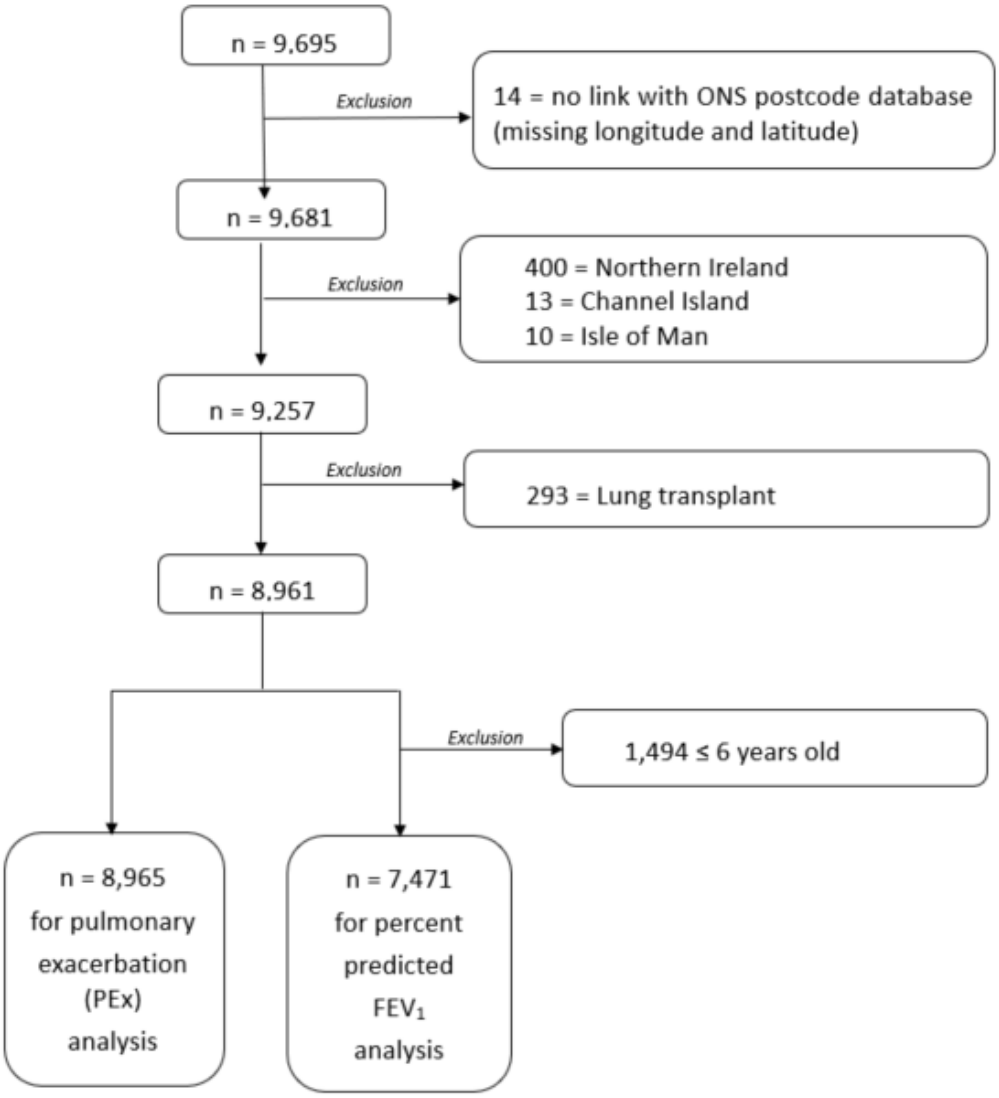


Figure 2.3: Flow diagram of the inclusion and exclusion criteria used for pwCF in GB in the 2016 UK CF Registry database

The summary statistics for $ppFEV_1$, #IVdays, and air pollutants (NO_2 and $PM_{2.5}$) are shown in ***Table 2.4***. The average $ppFEV_1$ was 70.6% (SD: 24.1), ranging from 11.0% to 145.8%. In our study, the average #IVdays for PEx was 16.7. (SD: 28.5) per year. The average NO_2 concentration among pwCF was 19.0 $\mu g/m^3$ (SD: 6.9), with a range of 1.6 to 71.2 $\mu g/m^3$ and the average $PM_{2.5}$ concentration was 9.4 $\mu g/m^3$ (SD: 1.3), with a range of 3.2 to 14.0 $\mu g/m^3$.

Table 2.4: Summary statistics of the outcome variable and air pollutants for pwCF with a residential postcode in Great Britain in 2016 on the UK CF Registry

Characteristics	n	mean	SD	Median	Range
$ppFEV_1$	7,147	70.6	24.1	73.4	11.0, 145.8
#IVdays	8,965	16.7	28.5	0.0	0, 329
NO_2 ($\mu g/m^3$)	8,965	19.0	6.9	18.6	1.6, 71.2
$PM_{2.5}$ ($\mu g/m^3$)	8,965	9.4	1.3	9.7	3.2, 14.0

$ppFEV_1$ = percent predicted FEV_1; #IVdays = Number of days on IV antibiotics for pulmonary exacerbation (PEx); SD = standard deviation

2.3.1.1 Demographic statistics of pwCF by health outcome

Table 2.5 compares the demographic characteristics of pwCF by $ppFEV_1$, #IVdays and air pollutants (NO_2 and $PM_{2.5}$). In my $ppFEV_1$ analysis group, 2,268 (30.4%) were children aged 7-16; 4,008 pwCF (53.6%) were male; 3,705 (49.6%) had homozygous *F508del*; 3,694 (49.4%) had at least one positive *P. aeruginosa* culture; and over 80% (n= 6,093) lived in rural areas. In total, 6,355 individuals (85.1%) had a residential postcode in England.

2.3.1.2 By lung function

The overall mean $ppFEV_1$ was lower in adults than in children (63.9% (SD: 23.7) *vs* 85.8% (SD: 17.2), p<0.001). There was no statistically significant difference in mean $ppFEV_1$ between males and females (70.8% *vs* 70.3%, p=0.323) (***Table 2*.5**). In terms of *CFTR* genotypes, the homozygous *F508del* (68.5%; SD: 24.0) and *F508del*/MF (68.5%; SD: 24.1) pwCF had the lowest mean $ppFEV_1$, while the RF/other group had the highest mean $ppFEV_1$ (78.7%; SD: 22.9). pwCF who had a positive *P. aeruginosa* culture had a lower mean $ppFEV_1$ (62.1% *vs* 79.0%, p<0.001) than those who were *P. aeruginosa* negative. Similarly, pwCF who lived in the most deprived (Townsend index) areas and urban areas had a lower mean $ppFEV_1$ (***Table 2*.5**). The mean $ppFEV_1$ of individuals in England was lower than that of people in Wales (70.4% *vs* 72%, p = 0.350).

2.3.1.3 By #IVdays for PEx

In young children (< 6 years), the mean #IVdays was 5.3 (SD: 11.2) days per year, while in adults (≥ 16 years), the mean was 21.7 (SD: 33.2) (***Table 2.5***). Females had a higher number

of #IVdays than males (19.6 *vs* 14.2, p<0.001). Individuals with homozygous *F508del* mutation had the highest (19.8, SD: 30.2) #IVdays, whereas those with RF/other had the lowest (6.5, SD: 17.4). Individuals with a positive *P. aeruginosa* culture had almost three times more days on IVs than those with a negative culture (26.0 *vs* 9.2, p<0.001). Finally, those living in the most deprived (Townsend index) areas (20.3 *vs* 14.4, p<0.001), and those residing in England had more #IVdays per year compared to those living in Wales (17.1 vs 14.7, p=0.002). The mean environmental exposure for NO_2 and $PM_{2.5}$ for each set of demographic characteristics is also shown in ***Table 2.5***.

Table 2.5: Demographic characteristics of pwCF with residential postcodes of GB

Characteristics	Health outcome						Environmental exposure	
	ppFEV$_1$ (n = 7,471)			**#IVdays (n = 8,965)**			**NO$_2$ (n = 8,965)**	**PM$_{2.5}$ (n = 8,965)**
	n	**%**	Mean (SD) - Range	**n**	**%**	Mean (SD) - Range	Mean (SD) - Range	Mean (SD) - Range
Age								
Children < 6	-*		-	1,494	16.7	5.3 (11.2) – 0, 94	19.1 (6.7) – 1.6, 68.4	9.5 (1.2) – 3.2, 14.0
6 - 15	2,268	30.4	85.8 (17.2) – 16.9, 145.8	2,268	25.3	12.7 (21.2) – 0, 180	18.5 (6.6) – 1.7, 54.2	9.4 (1.3) – 3.2, 13.4
Adults ≥ 16	5,203	69.6	63.9 (23.7) – 11.0, 132.1	5,203	58.0	21.7 (33.2) – 0, 329	19.2 (7.0) – 1.7, 71.2	9.4 (1.3) – 3.2, 13.9
Sex								
Male	4,008	53.6	70.8 (24.2) - 11.0, 132.9	4,790	53.4	14.2 (25.5) – 0, 279	19.0 (6.7) – 1.7, 68.4	9.4 (1.2) – 3.2, 14.0
Female	3,463	46.4	70.3 (24.0) – 15.4, 145.8	4,175	46.5	19.6 (31.4) – 0, 329	19.1 (7.1) – 1.6, 71.2	9.4 (1.3) – 3.2, 13.9
Genotype category								
homozygous F508del	3,705	49.6	68.5 (24.0) – 11.0, 145.8	4,456	49.7	19.8 (30.2) – 0, 318	18.9 (6.6) – 1.7, 54.2	9.4 (1.2) – 3.2, 13.5
F508del/MF	1,460	19.5	68.5 (24.1) – 11.1, 129.5	1,757	19.6	18.7 (31.2) – 0, 329	18.6 (6.6) – 2.3, 52.8	9.4 (1.2) – 3.2, 13.1
Residual (RF)/other	801	10.7	78.7 (22.9) – 15.9, 130.7	980	10.9	6.5 (17.4) – 0, 182	18.8 (7.2) – 1.6, 68.4	9.2 (1.4) – 3.2, 14.0
F508del/other	531	7.1	73.3 (24.5) – 14.3, 129.4	631	7.0	12.9 (25.0) – 0, 185	19.0 (7.2) – 3.3, 71.2	9.5 (1.3) – 3.3,13.5
Gating/any**	468	6.3	75.4 (24.0) – 16.3, 132.1	555	6.2	8.8 (17.8) – 0, 132	18.8 (6.8) – 2.2, 54.2	9.3 (1.3) – 3.2, 13.1
MF/MF	243	3.5	69.0 (21.8) – 15.6, 121.4	291	3.2	18.5 (27.2) – 0, 208	21.1(8.3) – 1.7, 49.9	9.7 (1.4) – 3.2, 13.9
Other/Other	139	1.9	74.4 (23.9) – 20.1, 115.3	157	1.8	13.2 (32.5) – 0, 283	22.2 (7.5) – 3.2, 49.1	9.9 (1.0) – 4.2, 11.8
MF/other	124	1.7	74.6 (23.8) – 19.3, 118.7	138	1.5	12.6 (22.4) – 0, 142	21.4 (8.9) – 5.6, 55.3	9.6 (1.4) – 3.7, 13.9
P. aeruginosa culture								
Negative	3,777	50.6	79.0 (21.8) – 11.0, 145.8	4,863	54.8	9.2 (20.5) – 0, 318	18.7 (6.9) – 1.6, 71.2	9.4 (1.3) – 3.2, 14.0
Positive	3,694	49.4	62.1 (23.3) – 11.1, 132.1	4,012	45.2	26.0 (33.9) – 0, 329	19.4 (6.9) – 1.7, 52.8	9.5 (1.2) - 3.2, 13.9
Deprivation (Townsend index)								
Q1 (Least deprived)	1,567	20.1	72.7 (23.6) – 13.6, 129.5	1,855	20.7	14.4 (25.4) – 0, 318	15.9 (5.0) – 1.7, 34.4	9.1 (1.2) – 3.2, 11.8
Q2	1,556	20.1	70.4 (24.8) – 11.0,145.8	1,875	20.9	16.0 (28.2) – 0, 315	16.9 (5.9) – 1.6, 47.8	9.3 (1.2) – 3.2, 12.6
Q3	1,565	20.1	70.4 (24.1) – 11.1, 132.6	1,863	20.8	16.1 (27.4) – 0, 329	18.7 (6.3) – 2.0, 47.0	9.4 (1.2) – 3.2, 13.0
Q4	1,481	19.8	69.9 (23.6) – 15.4, 132.9	1,784	19.9	17.3 (29.4) – 0, 287	20.5 (6.7) – 4.0, 68.4	9.5 (1.2) – 3.7, 14.0
Q5 (most deprived)	1,302	17.4	69.1 (24.2) – 12.1, 133.3	1,588	17.7	20.3 (32.2) – 0, 238	23.9 (7.7) – 4.5, 71.2	9.9 (1.2) – 3.2, 13.9

Rural and Urban								
Rural	1,378	18.4	71.5 (24.3) – 13.6, 127.3	1,651	18.4	16.1 (26.8) – 0, 189	11.9 (4.1) – 1.6, 34.9	8.7 (1.5) – 3.2, 11.5
Urban	6,093	81.6	70.3 (24.1) – 11.0, 145.8	7,314	81.6	16.8 (28.9) – 0, 329	20.6 (6.4) – 3.2, 71.2	9.6 (1.1) – 3.4, 14.0
Country								
England	6,355	85.1	70.4 (24.1) – 11.0, 145.8	7,675	85.6	17.1 (28.7) – 0, 329	19.8 (6.8) – 2.9, 68.4	9.7 (1.0) – 4.1, 14.0
Scotland	686	9.2	71.1 (24.8) – 15.5, 125.8	796	8.9	13.9 (25.7) – 0, 182	15.1 (6.1) – 1.6, 71.2	7.3 (1.2) – 3.2, 13.5
Wales	430	5.7	72.0 (23.2) – 16.3, 126.0	494	5.5	14.7 (25.7) – 0, 157	13.5 (4.8) – 3.2, 34.1	8.3 (5.2) – 3.3, 11.2

$ppFEV_1$ = percent predicted FEV_1; #IVdays = Number of days on IV antibiotics for pulmonary exacerbation (PEx); *FEV_1 not reliable under 6 years of age; GB = Great Britain

MF = Minimal function; RF = Residual function; Other = Either the mutation was not identified, or it was identified but classification into one of the other categories was not possible.

**Except for RF mutation; *P. aeruginosa= Pseudomonas Aeruginosa*

2.3.2 Correlation between outcome variables and air pollutants

The correlation matrix was generated between key outcome variables ($ppFEV_1$, #IVdays) and air pollutants (NO_2 and $PM_{2.5}$). ***Table 2.6*** shows the correlation coefficients between NO_2, $PM_{2.5}$, $ppFEV_1$ and #IVdays. The correlation coefficient between NO_2 and $ppFEV_1$ was small ($p > 0.05$). I found a significant correlation between $PM_{2.5}$ and #IVdays ($p < 0.05$). NO_2 and $PM_{2.5}$ were highly correlated with each other ($p < 0.001$). Therefore, each pollutant was evaluated separately.

Table 2.6: Correlation between air pollutants and key outcome variables among pwCF with a residential postcode of GB

	NO_2	**$PM_{2.5}$**	**$ppFEV_1$**	**# IV days**
NO_2	1.00			
$PM_{2.5}$	0.66****	1.00		
$ppFEV_1$	-0.02	0.01	1.00	
# IV days	0.01	0.03**	-0.48****	1.00

p <0 .0001, "****", p < 0.001, "***", p < 0.01, "**", p < 0.05, "*"

The relationship between $ppFEV_1$, #IVdays and NO_2 and $PM_{2.5}$ in children and adults with CF are shown in scatter plots in ***Figures 2.4 and 2.5.*** A series of box plots were generated for the two health outcomes considered with quintiles of NO_2 and $PM_{2.5}$ to explore the dispersion of the data. All of these graphs can be seen in ***Appendix 7.1-7.16***: Chapter II.

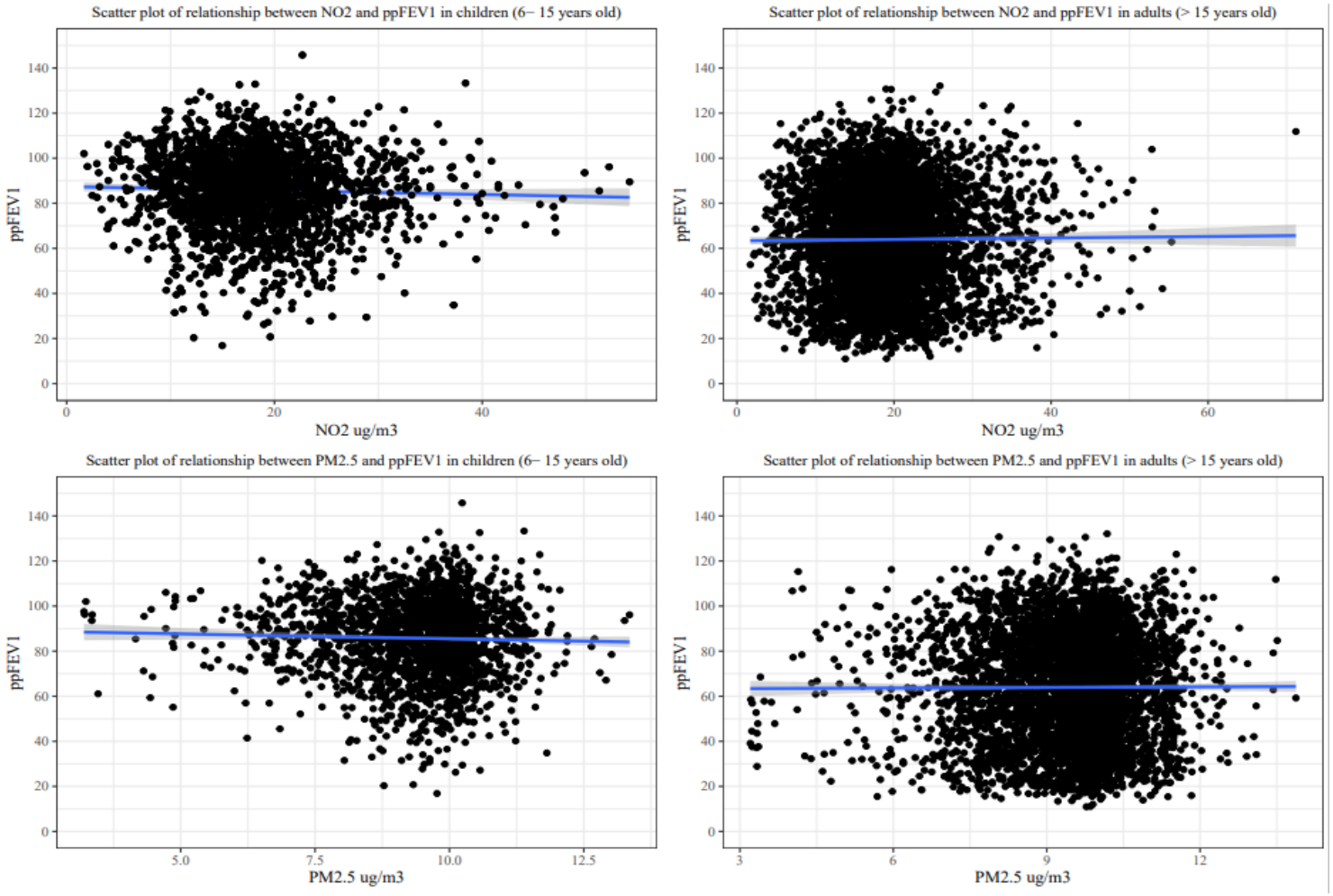


Figure 2.4: Scatter plot of the relationship between ppFEV$_1$ and NO_2 (top) and $PM_{2.5}$ (bottom) in µg/m^3 in children (left) and adults (right) with CF in GB

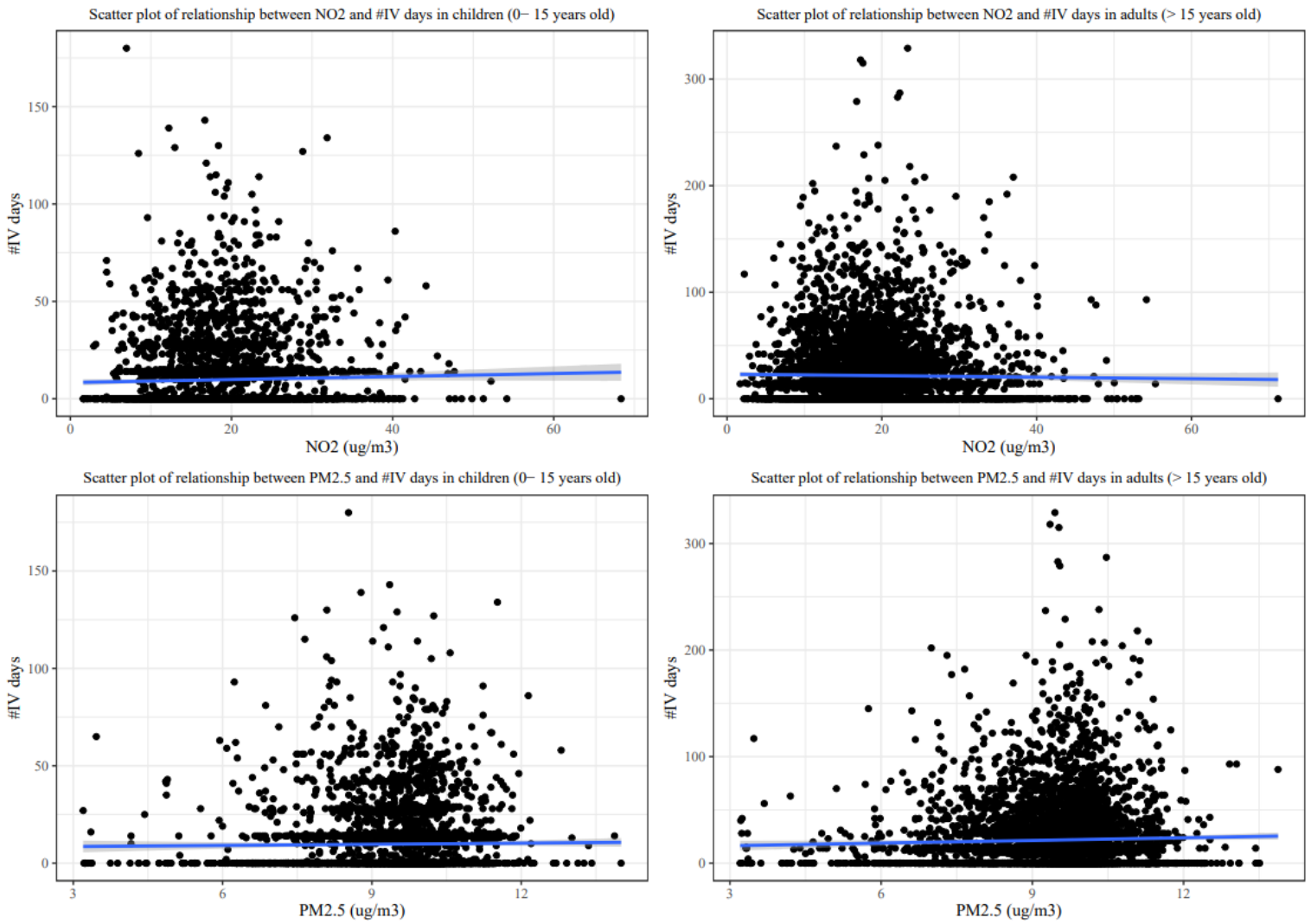


Figure 2.5: Scatter plot of the relationship between #IVdays and NO_2 and $PM_{2.5}$ in µg/m^3 in children and adults with CF in GB

2.3.3 Association between ppFEV$_1$ and covariates in children and adults

The findings of my univariate regression analysis of lung function and covariates in children and adults are shown in ***Table 2.7***. In children with CF, there was a statistically significant association between ppFEV$_1$ and group, genotype groups, *P. aeruginosa* cultures, deprivation, rurality, and country. I found a lower ppFEV$_1$ (-1.74%, 95%CI: -3.19; -0.30, p=0.018) in female, children with positive *P. aeruginosa* (-5.35%, 95%CI: -6.93; -3.76, p<0.001), children in deprivation quintile Q4 (-3.47, 95%CI: -5.71; -1.23, p=0.005) and living in urban areas (-1.56, 95%CI: -3.38; -0.26, p=0.094). Children with a residual (8.60%, 95%CI; 5.98 and 11.20%) or gating (5.29%, 95%CI: 2.06; 8.52%) genotype category had a higher ppFEV$_1$ compared to homozygous F508del children (p<0.001). Similarly, children with CF in Scotland had a ppFEV$_1$ 3.67% higher (95%CI; 1.13, 6.20, p= 0.017) than those living in England.

Similar results are described for adults in ***Table 2.7***. I found that adults with positive *P. aeruginosa* culture had lower ppFEV$_1$ 14.8% (95%CI: -16.00; -13.50, p<0.001) than those who were *P. aeruginosa* negative. In comparison to adults in the least deprived area (Townsend index), those in the most deprived area had a lower ppFEV$_1$ of 4.37% (95% CI: -6.53; -2.21, p = 0.003). However, adults with Residual and Gating mutations had higher ppFEV$_1$ of 13.1% (95%CI: 11.0; 15.2%) and 9.53% (95% CI: 6.89; 12.2%), respectively.

Table 2.7: Univariate linear regressions between $ppFEV_1$ and characteristics of children and adults with CF in the UK CF Registry in GB

	$ppFEV_1$					
	Children			**adults**		
Characteristic	**Coefficient**	**95% CI**	**p-value**	**Coefficient**	**95% CI**	**p-value**
Sex			**0.018**			0.060
Male	Ref	-		Ref	-	
Female	-1.74	-3.19, -0.30		-1.27	-2.59, 0.06	
Genotype			**<0.001**			**<0.001**
homozygous F508del	Ref	-		Ref	-	
F508del/MF	0.92	-0.94, 2.78		-0.08	-1.82, 1.65	
Residual/any	8.60	5.98, 11.20		13.1	11.0, 15.2	
F508del/other	4.64	1.38, 7.89		7.76	5.27, 10.2	
Gating/any**	5.29	2.06, 8.52		9.53	6.89, 12.2	
MF/MF	-2.63	-6.12, 0.85		-0.99	-4.97, 2.99	
Other/Other	5.27	-2.42, 13.0		10.6	5.90, 15.3	
MF/other	6.46	0.40, 12.5		7.94	2.98, 12.9	
***P. aeruginosa* culture**			**<0.001**			**<0.001**
negative	Ref			Ref		
positive	-5.35	-6.93, -3.76		-14.80	-16.0, -13.5	
Deprivation (Townsend index)			**0.005**			**0.003**
Q1 (Least deprived)	Ref	-		Ref	-	
Q2	0.63	-1.59, 2.85		-2.58	-4.61, -0.55	
Q3	-0.88	-3.09, 1.32		-2.23	-4.27, -0.20	
Q4	-3.47	-5.71, -1.23		-1.80	-3.86, -0.26	
Q5 (most deprived)	-1.66	-3.96, 0.63		-4.37	-6.53, -2.21	
Rural and Urban			0.094			0.712
Rural	Ref	-		Ref	-	
Urban	-1.56	-3.38, 0.26		-0.32	-2.04, 1.39	
Country			**0.017**			0.483
England	Ref	-		Ref	-	
Scotland	3.67	1.13, 6.20		-0.42	-2.70, 1.86	
Wales	-0.23	-3.21, 2.75		1.64	-1.23, 4.51	-0.23

$ppFEV_1$ = percent predicted FEV1 MF = Minimal function; RF = Residual function; Other = Either the mutation was not identified, or it was identified but classification into one of the other categories was not possible. **Except for RF mutation.

2.3.4 Lung function and air pollutants (NO_2 and $PM_{2.5}$)

2.3.4.1 Lung function and NO_2 in children

In the univariate linear regression analysis, $ppFEV_1$ was lower in children across all NO_2 quintiles compared to the lowest quintile (Q1) but it did not reach statistical significance. $ppFEV_1$ was 2.18% (95%CI; -4.45; 0.09) and 1.88% (95%CI: 4.19; 0.43) lower in the Q4 and Q5 quintiles of NO_2, respectively, in unadjusted linear regression analysis. This difference in $ppFEV_1$ became smaller and the association remained not significant –0.98 (95%CI: -3.05; 1.64) and -0.70% (95%CI: -3.05; 1.64) after adjusting for age, sex, BMI percentile, genotype, *P. aeruginosa*, deprivation and country of residence in the adjusted multiple linear regression analysis *(**Figure 2.6**).*

In a sensitivity analysis of a sub-group of children who did not change their residential postcode during the five years preceding their annual review (2012 - 2016), the difference in $ppFEV_1$ became greater across all NO_2 quintiles. The mean $ppFEV_1$ was -1.52% (95%CI: -4.24; 1.19) and -0.88% (95%CI: -3.76; 2.00%) in Q4 and Q5 quintile of NO_2, respectively *(**Figure 2.6**)*.

2.3.4.2 Lung function and $PM_{2.5}$ in children

There was a significant association between lung function and $PM_{2.5}$. I found $ppFEV_1$ was -3.76% (95%CI; -6.03, -1.50%) in Q2 of $PM_{2.5}$ compared to the lowest quintile (Q1) (***Figure 2.6***). This association remained statistically significant (-3.02, 95%CI; -5.15, -0.93, p=0.004) in Q2 of $PM_{2.5}$ after adjusting for age, sex, BMI percentile, genotype, *P. aeruginosa*, and deprivation (Townsend index) in the multiple linear regression analysis. This association was attenuated in a sensitivity analysis of a subgroup of children who did not change their residential postcode between 2012 and 2016 (***Figure 2.6***).

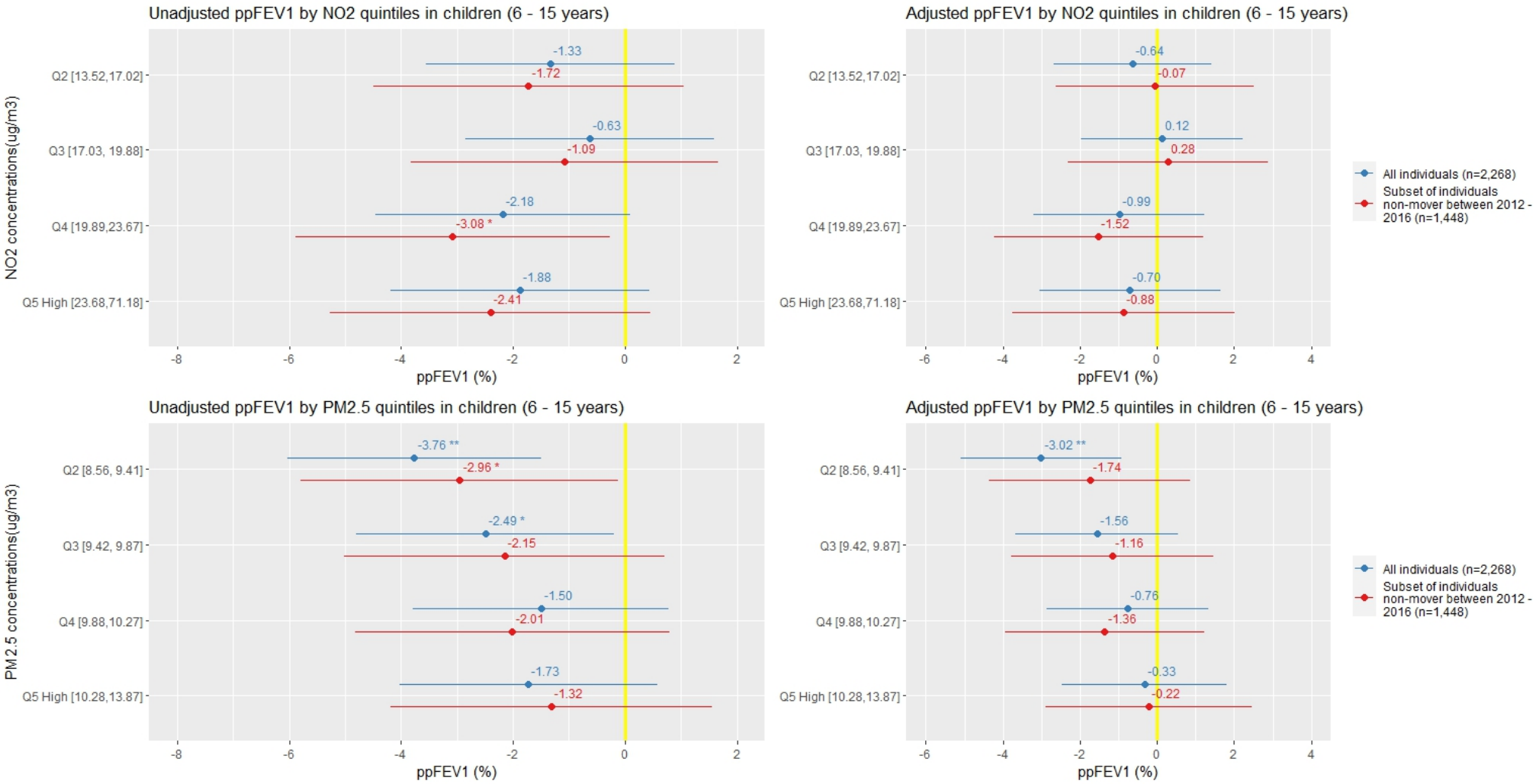


Figure 2.6: Linear regressions between air pollutants (NO_2 and $PM_{2.5}$) and lung function in children (age 6 – 15 years) with CF in the UK CF Registry in GB

NO_2 model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation, country and sex
$PM_{2.5}$ model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation and sex

2.3.4.3 Lung function and NO_2 in adults

The linear regression analysis of lung function and air pollutants in adults is shown in ***Figure 2.7***. There was no significant association between $ppFEV_1$ and NO_2 in all adults and in a subgroup of individuals who did not change their residential postcode between 2012 and 2016.

In adults, the mean $ppFEV_1$ in the Q5 quintile of NO_2 was 0.38% (95%CI; -1.71, 2.46%) and 2.11% (95%CI; 0.07, 4.16%) in unadjusted and adjusted linear regression analyses, respectively. Similarly, in a sensitivity analysis, there was no statistically significant association between lung function and NO_2 in a sample of adults who did not change address between 2012 and 2016 (1.08%, 95%CI; -1.76, 3.92, p=0.454 in Q5 quintile of NO_2) (***Figure 2.7***).

2.3.4.4 Lung function and $PM_{2.5}$ in adults

Figure 2.7 summarises the findings of the linear regression analysis for the association between $PM_{2.5}$ and $ppFEV_1$ in adults. All associations between $PM_{2.5}$ and $ppFEV_1$ in adults, both unadjusted and adjusted, show a negative beta coefficient, although only one is statistically significant (2.96%, 95%CI; -5.79, -0.12%, p=0.041). This association attenuated (-1.80, 95%CI; - 4.42, 0.80, p=0.173) after for *P. aeruginosa*, absolute BMI, age, genotype, deprivation and sex in the multiple linear regression analysis. Results for the subset of adults who did not move during the 5 years preceding their 2016 annual review are less clear (***Figure 2.7***).

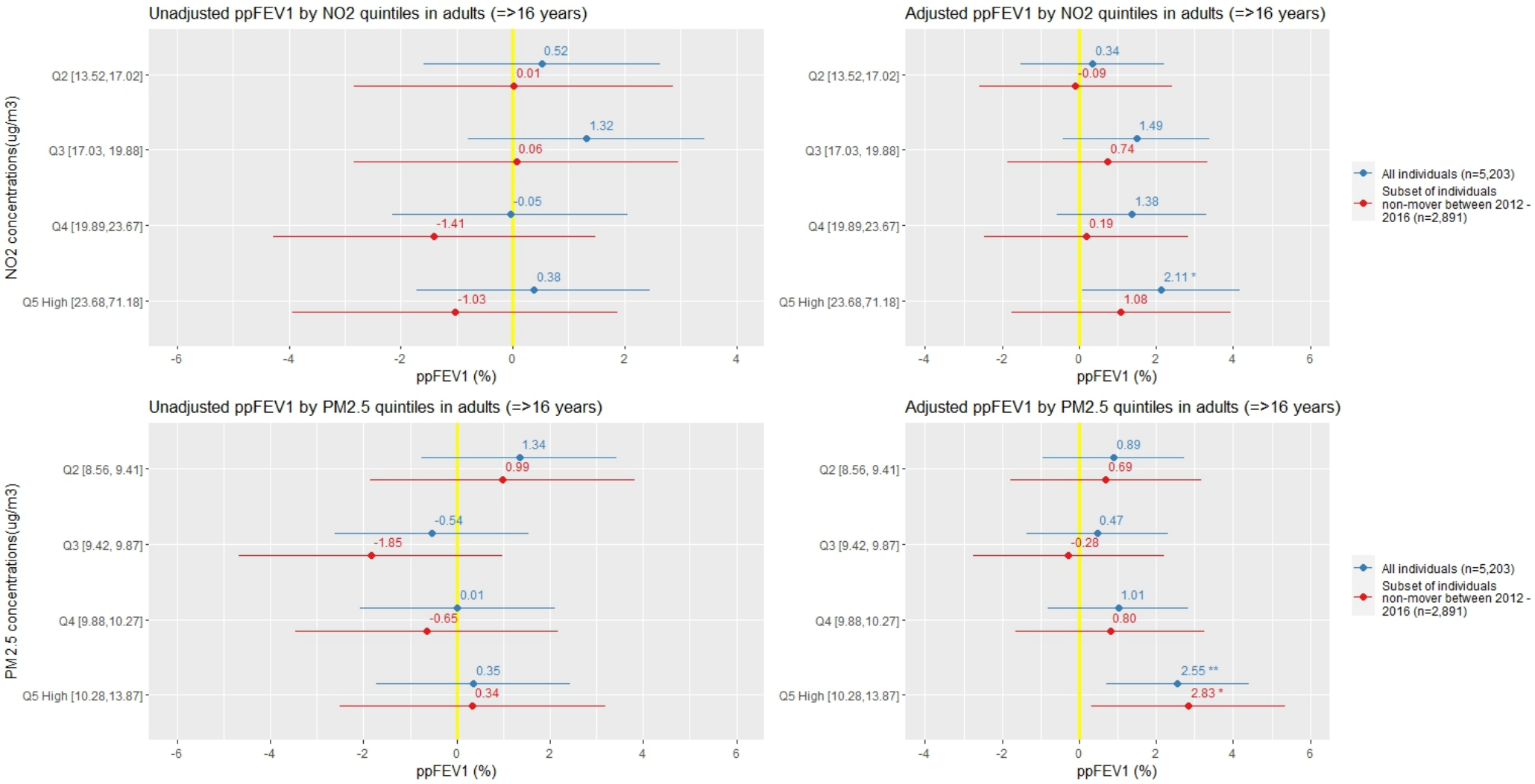


Figure 2.7: Linear regression between air pollutants (NO_2 and $PM_{2.5}$) and lung function in adults (age ≥ 16) with CF in the UK CF Registry in GB

NO_2 model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation, country and sex
$PM_{2.5}$ model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation and sex

2.3.5 #IVdays and air pollutants (NO_2 and $PM_{2.5}$)

2.3.5.1 #IVdays and NO_2 in children

Figure 2.8 displays the quintiles NO_2 of the ZINB model's regression coefficients for children with CF. In the count model, I found that children had higher numbers of IV days in all NO_2 quintiles except Q5. The rate of #IVdays was 0.98 (95%CI; 0.85, 1.11, p=0.742) days per year. In the zero-inflated unadjusted regression model, the odds of having zero IV days were lower in higher quintiles of $PM_{2.5}$ (0.79, 95%CI: 0.64, 0.98, p=0.032). After adjusting for age, sex, BMI percentile, lung function (ppFEV$_1$), *P. aeruginosa*, genotype, and deprivation, this statistical association was attenuated (1.10, 95%CI; 0.82, 1.49, p=0.531).

2.3.5.2 #IVdays and NO_2 in adults

The regression coefficients for the quintiles NO_2 in the ZINB model in the adult group are shown in ***Figure 2.9***. Overall, adults with CF showed no statistically significant association between NO_2 and # IV days. In the count model, I found 0.96 (95%CI; 0.88, 1.05, p=0.386) and 0.92 (95%CI; 0.85, 1.01, p=0.067) in the rate of #IVdays per year before and after adjustment for age, sex, BMI, genotype, *P. aeruginosa* and deprivation in Q5 quintile of $PM_{2.5}$.

2.3.5.3 #IVdays and $PM_{2.5}$ in children

There was an effect of $PM_{2.5}$ on #IVdays among children in our count model of ZINB analysis ***Figure 2.10***. I found that the rate of #IVdays per year in Q5 quintile $PM_{2.5}$ was 0.82 (95%CI: 0.72, 0.93, p=0.002) lower in the count model. In our zero-inflation model, the likelihood of zero #IVdays was lower for higher quintiles of $PM_{2.5}$ compared to Q1's lower quartile. In our unadjusted zero-inflation model, the odds of experiencing zero #IVdays was 0.72 (95% CI: 0.58, 0.89, p=0.002) days per year in the Q5 quintile $PM_{2.5}$. This association was no longer statistically significant in our adjusted model (0.92, 95%CI;0.70, 1.22, p=0.559).

2.3.5.4 #IVdays and $PM_{2.5}$ in adults

Figure 2.11 displays the ZINB regression coefficients for the $PM_{2.5}$ quintiles and #IVdays for adults with CF. There was no linear trend across the $PM_{2.5}$ quintiles and #IVdays, however, I found a statistically significant association between $PM_{2.5}$ and #IVdays in adults with CF. In the count model, I identified an increase of 1.11 (95%CI: 1.01-1.21, p=0.036) and 1.08 (95%CI:

0.89,1.17, p=0.087) in #IVdays per year in Q5 quintile $PM_{2.5}$ before and after adjustment for age, sex, absolute BMI, lung function ($ppFEV_1$), *P. aeruginosa*, genotype and deprivation among all adults with CF (***Figure 2.12***). The odds of experiencing zero IV days were lower in higher quintiles of $PM_{2.5}$ in the zero-inflated model, albeit this difference was not statistically significant in the Q5 quintile of $PM_{2.5}$ (1.07, 95%CI; 0.87, 1.32, p=0.506).

When comparing the sensitivity analysis of the subgroup of pwCF who did not change their residential postcode to the complete data set (all adults) study, I found the same strength of relationship and coefficient of $PM_{2.5}$ quintiles.

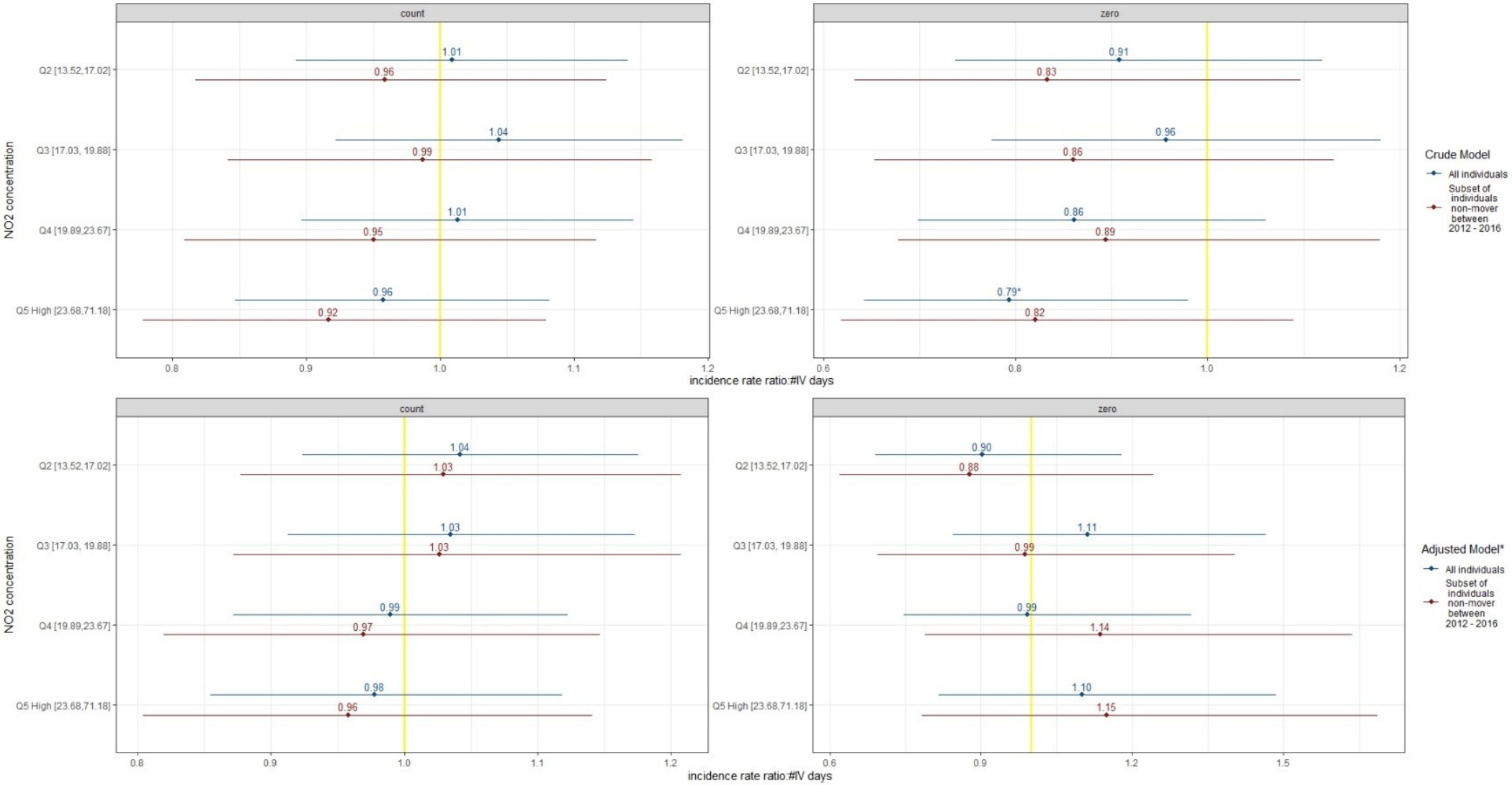


Figure 2.8: Zero-inflated negative binomial regression analysis between NO_2 and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB

NO_2 model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation, country and sex

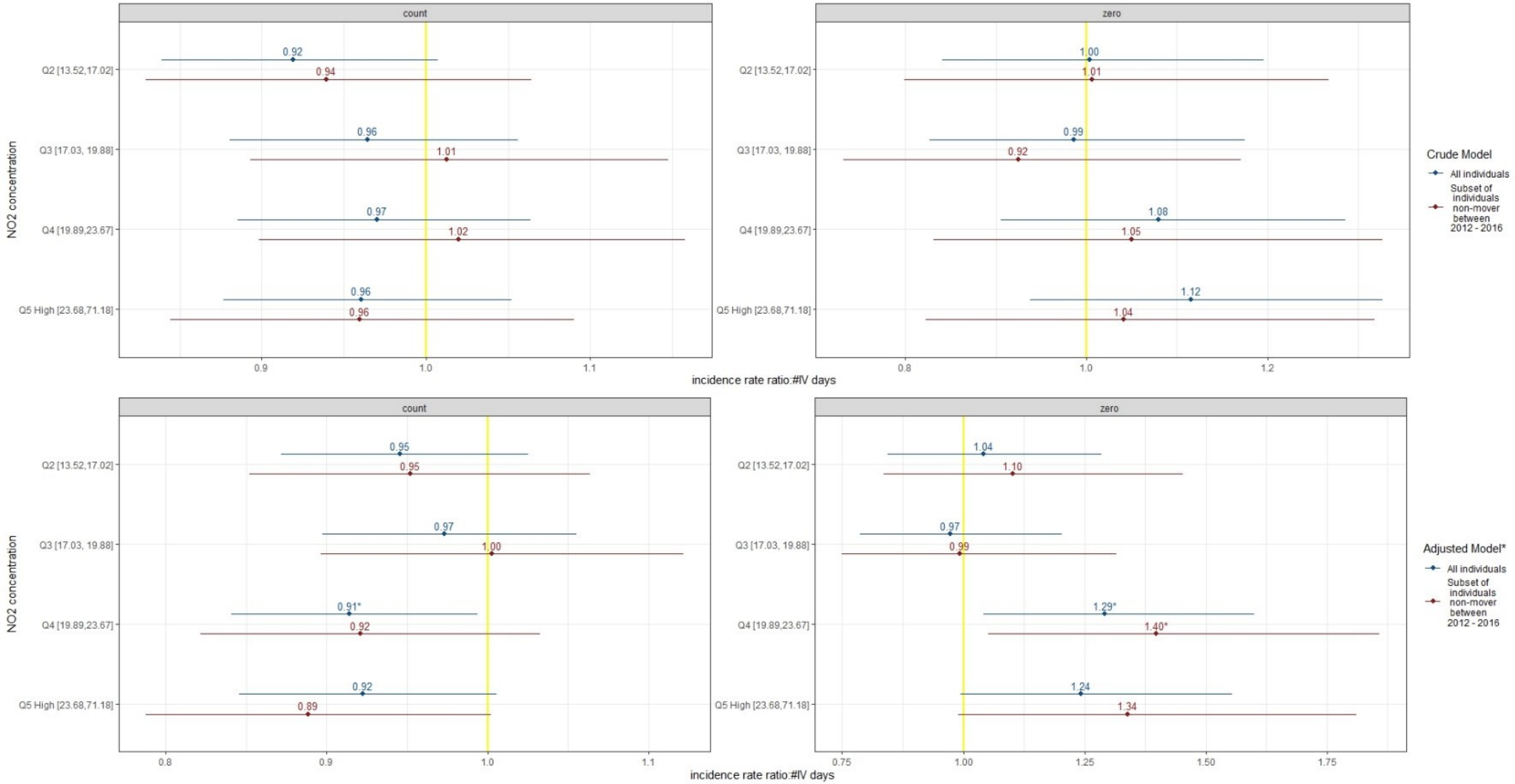


Figure 2.9: Zero-inflated negative binomial regression analysis between NO_2 and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB

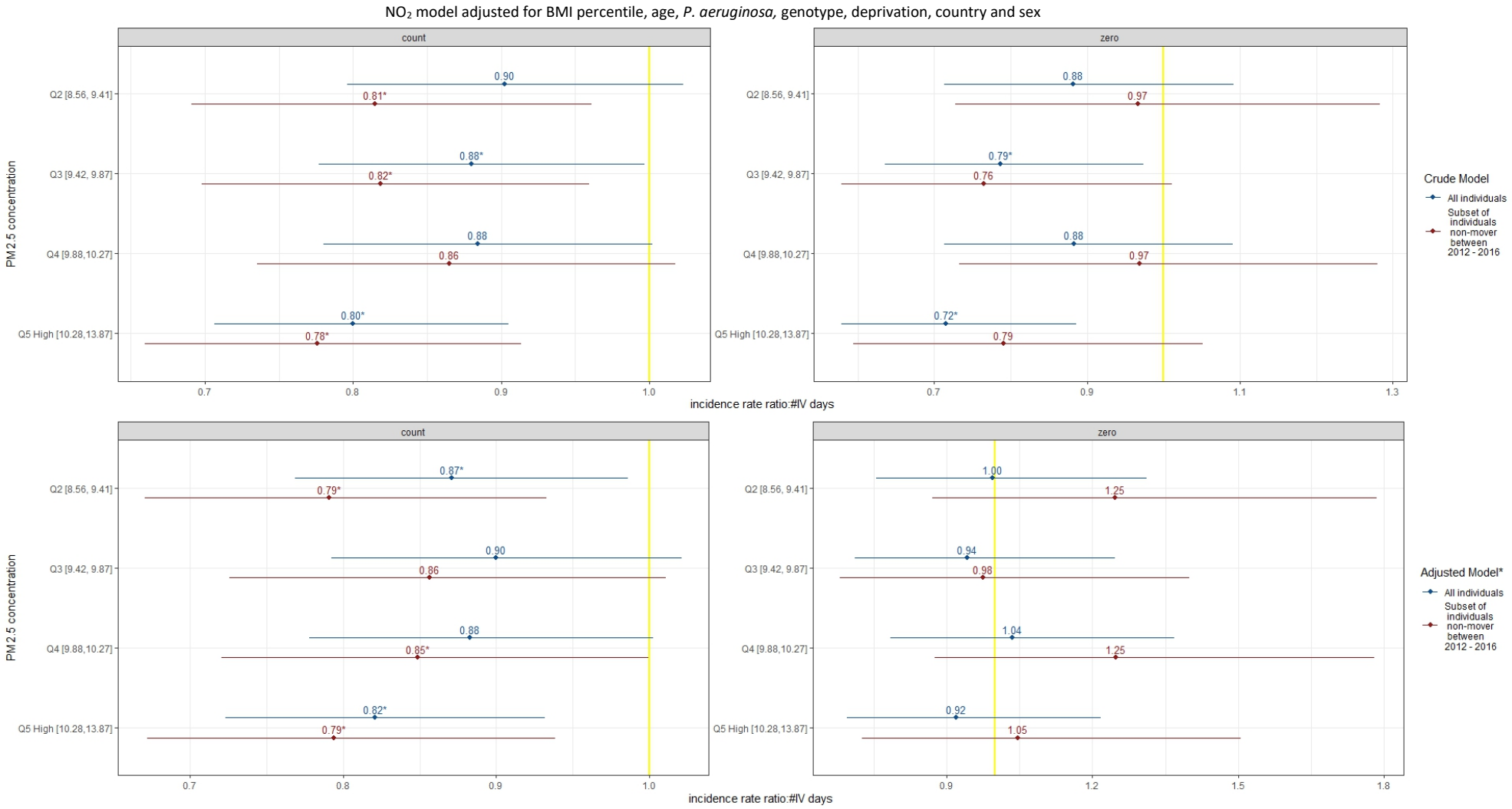


Figure 2.10: ZINB regression analysis between $PM_{2.5}$ and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB

$PM_{2.5}$ model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation and sex

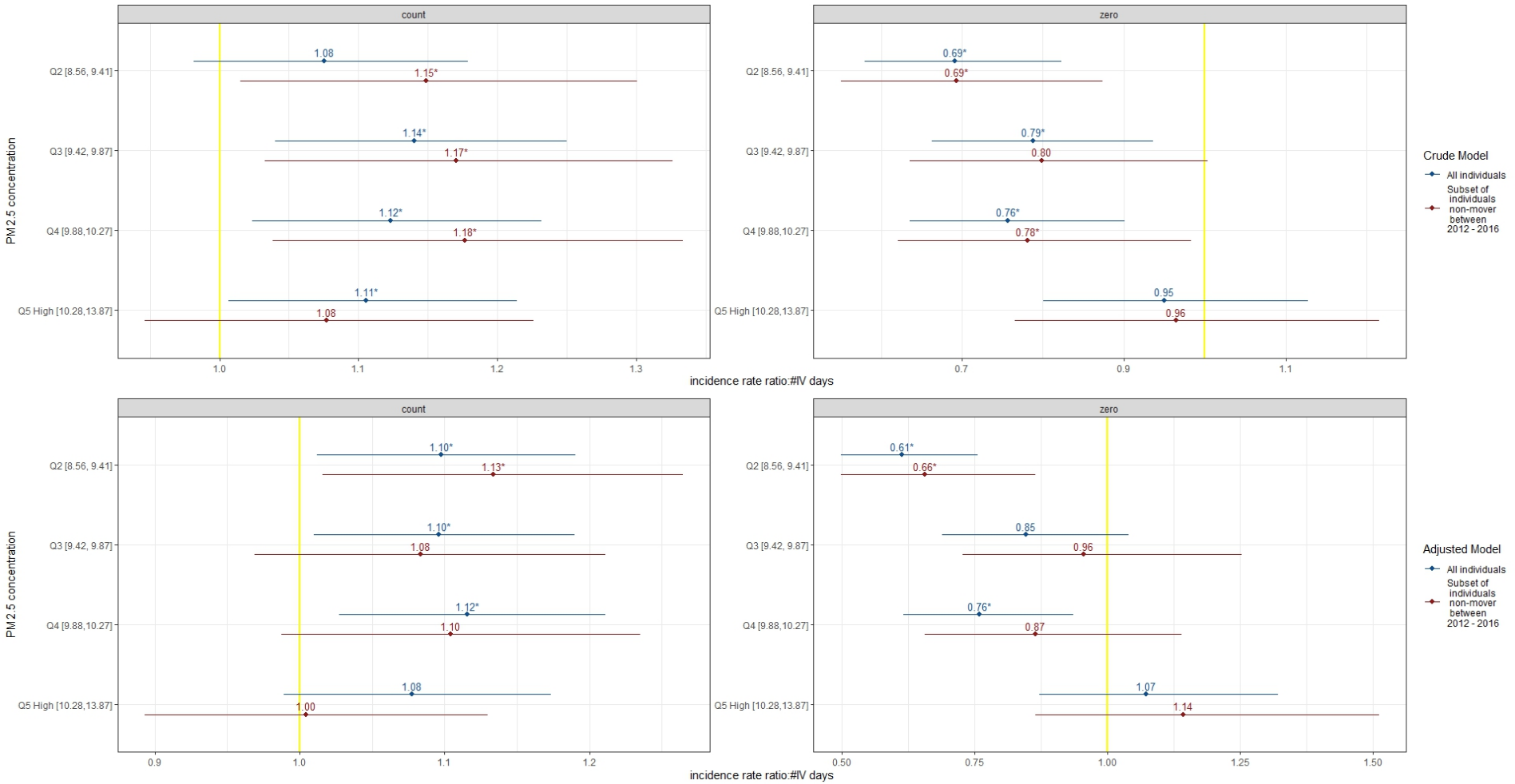


Figure 2.11: Zero-inflated negative binomial regression analysis between $PM_{2.5}$ and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB

$PM_{2.5}$ model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation and sex

2.3.6 Sensitivity analyses

Table 2.8 summarises the distribution of #IVday summary statistics overall, excluding days more than 180 and 120. 4,755 out of the 8,965 pwCF, or more than half, had no #IVdays. The range of #IVdays per year was between 0 and 320, with a mean of 16.7 days and a variance of 814.3.

Table 2.8: Distribution of the summary statistics of #IV day for PEx in pwCF

Characteristics	**All patients**	**Excluding above 180 # IV days (n=29)**	**Excluding above 120 # IV days (n=130)**
n	8965	8936	8835
Zero n (%)	4,755 (53.0%)	4,755 (53.2%)	4,755 (53.8%)
Mean (IV days)	16.7	16.04	14.6
Median (IV days)	0.0	0.0	0.0
Variance (IV days)	814.3	674.4	503.1
Range (IV days)	0, 320	0, 180	0, 120

Six patients had extremely low ppFEV$_1$ (below 15%) at their annual visit in 2016. These patients' ppFEV$_1$ were cross-checked in the last 5 years (2012-2016) annual visits and it showed a consistent low trend during this period in all patients except one, where there was a sudden step drop (24% in 2016 and 11% in 2015). This could be possible due to the end state of their lungs.

Almost 60% of pwCF did not change their residence postcode in the five years before the annual visit in 2016, while 88.8% of pwCF had lived in two different postcodes.

In my sensitivity analysis of the subgroup of children with CF who did not change their residential postcode between 2012 and 2016, the impact of NO_2 on ppFEV$_1$ was stronger across all NO_2 quintiles in the unadjusted model compared to all children with CF ***(Figure 2.6 & Figure 7.1)***. This association was attenuated in the adjusted model. In stratified genotype analyses, there was no clear signal of association between NO_2, $PM_{2.5}$, and lung function in CF adults. (***Table 2.9***)

Table 2.9: Linear regression between air pollutants (NO_2 and $PM_{2.5}$) and lung function in adults (age ≥ 16) by mutation (severe vs mild) with CF in GB

	$ppFEV_1$											
	Severe genotype n = 5,876						**Mild genotype n = 1,595**					
	Unadjusted			**Adjusted**			**Unadjusted**			**Adjusted**		
Characteristic	**β**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**	**B**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (13.52, 17.02)	1.64	-0.68, 3.97	0.166	0.98	-1.09, 3.05	0.354	-3.39	-7.99 1.20	0.148	-1.71	-5.74, 2.33	0.408
Q3 (17.03, 19.88)	0.85	-1.48, 3.19	0.474	0.73	-1.38, 2.84	0.495	2.46	-2.07, 7.00	0.287	4.08	0.05, 8.11	**0.047**
Q4 (19.89, 23.67)	0.15	-2.16, 2.46	0.901	1.18	-0.96, 3.32	0.279	0.02	-4.62, 4.67	0.992	2.73	-1.59, 7.06	0.216
Q5 – High (23.68, 71.18)	-0.10	-2.43, 2.23	0.933	2.03	-0.26, 4.33	0.083	-0.08	-4.43, 4.23	0.969	3.24	-1.02, 7.50	0.136
AIC		35037.5			34025			10465.0			10154.9	
$PM_{2.5}$ (2015 model) (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (8.56, 9.41)	2.04	-0.29, 4.36	0.086	1.18	- 1.23, 3.58	0.338	-0.90	-5.33, 3.53	0.690	-2.04	-6.71, 2.63	0.392
Q3 (9.42, 9.87)	0.36	-1.96, 2.67	0.761	0.61	-1.85, 3.07	0.629	-3.03	-7.51, 1.46	0.186	-2.81	-7.59, 1.97	0.249
Q4 (9.88, 10.27)	0.27	-2.03, 2.56	0.819	0.50	-1.98, 2.98	0.693	0.81	-3.77, 5.39	0.728	0.80	-4.16, 5.76	0.751
Q5 – High (10.28, 13.87)	0.73	-1.59, 3.05	0.536	2.53	-0.01, 5.07	0.051	-1.76	-6.06, 2.53	0.421	0.54	-4.27, 5.34	0.827
AIC		35037			34022.3			10468.3			10160.4	

CF = cystic fibrosis; $ppFEV_1$ = percent predicted FEV_1; AIC= Akaike information criterion; GB= Great Britain

NO_2 model adjusted for *P. aeruginosa*, absolute BMI, age, deprivation, country and sex (double check country in adults)

$PM_{2.5}$ model adjusted for *P. aeruginosa*, absolute BMI, age, deprivation and sex

Severe genotype= homozygous F508del, F508del/MF, Gating/any**, MF/MF (**Except RF mutation)

Mild genotype= Residual (RF)/other, F508del/other, Other/Other, MF/other

2.4 Discussion

To my knowledge, this is one of the largest studies conducted to investigate the relationship between air pollutants (NO_2 and $PM_{2.5}$) and disease severity (PEx and $ppFEV_1$) among pwCF, and the first one in the UK. This study of children and adults with CF used high-quality data from the UK CF Registry and a novel high-resolution continuous modelled surface for annual average concentrations of NO_2 and $PM_{2.5}$ in GB, while also accounting for age, sex, BMI, *P. aeruginosa*, genotype, deprivation, and country in the regression models. The main finding of this study was the association between $PM_{2.5}$ and #IVdays in adults with CF. Additionally, there was an association between $PM_{2.5}$ and $ppFEV_1$ in children with CF. However, there was no statistically significant association between NO_2 and $ppFEV_1$ or #IVdays in either children or adults with CF. Below, I discuss these findings in sequence, in light of existing evidence from the literature.

2.4.1 Lung function ($ppFEV_1$)

2.4.1.1 $ppFEV_1$ and NO_2

Lung function is an established and reliable marker of CF disease progression that is used to assess the clinical course and evaluate the effectiveness of treatment (Szczesniak *et al.*, 2017a). My study found only a marginal difference in $ppFEV_1$ across NO_2 quintiles in children with CF. In addition, there was no evidence of an association between $ppFEV_1$ and NO_2 in adults with CF. There is little epidemiological evidence of the effects of NO_2 on lung function in pwCF. To my knowledge, Goss *et al.* are the only other group to study this relationship; in keeping with the results of my investigation, they also did not find a relationship (Goss *et al.*, 2004).

Comparing the results from the current study's CF population to those from studies on other respiratory diseases may help researchers and epidemiologists understand more about how NO_2 affects CF. The evidence of an association between NO_2 and $ppFEV_1$ has been inconsistent in a non-CF population (Barone-Adesi *et al.*, 2015; Doiron *et al.*, 2019; Elbarbary *et al.*, 2020; Hou *et al.*, 2020; Kim & Song, 2017; Strassmann *et al.*, 2021; Xiao *et al.*, 2017). In healthy children in England, a cross-sectional study (CHASE) and meta-analysis exploring the long-term exposure of NO_2 to lung function did not find a significant association between NO_2 and lung function in 4,884 children with a mean age of 9.9 (*SD* = 0.4) years (Barone-Adesi *et al.*,

2015). NO_2 exposure was estimated at each participant's current home address at 20x20m resolution using the KCLurban dispersion model, developed as part of the TRAFFIC project (Beevers *et al.*, 2013). The CHASE study's main strengths were that it was one of the highest-powered studies that looked at how air pollution affects children's lungs over the long term; passive smoking exposure, which was measured using cotinine levels; and socioeconomic status, which was determined using the information on the parents' socioeconomic situations. The CHASE study included 4,884 healthy school-age children, whereas the present study had a slightly smaller sample size (3,762) but included children with CF. I used a high-resolution NO_2 map resolution of 25x25m in this project, which is comparable to the NO_2 map (20x20m) used in the CHASE study. When adjusting for potential confounding variables, the CHASE study performed better, especially when adjusting for passive smoking exposure, which was calculated using cotinine levels. Smoking data are limited in the UK CF Registry; as a result, they were not taken into account in the regression analysis. In the second part of the CHASE study, Barone-Adesi *et al.* did a meta-analysis, which included 13 studies published between 1990 and 2015 among children or adolescents aged between 0 and 18 years, reporting an 8 mL lower FEV_1 (95% CI: -14;-1 ml, p = 0.016) with every 10 µg/m^3 increase in NO_2 (Barone-Adesi *et al.*, 2015). Similarly, a COPD population-based UK Biobank study reported that for every 10 µg/m^3 increase in NO_2, 299,537 participants had a statistically significantly lower FEV_1 of -33.85 ml (95% CI: -36.34; -31.36) (Doiron *et al.*, 2019). The UK Biobank study and the current study both used LUR-based NO_2 estimates, making them comparable in terms of exposure assignment for NO_2. In contrast to the UK Biobank study, where the mean (SD) NO_2 concentration was 26.31 (7.49) µg/m^3 the mean NO_2 concentration in the current analysis was 19.0 (6.9) µg/m^3. These results may indicate that similar strong associations between $ppFEV_1$ and NO_2 in CF pollution would have been expected, but the associations between $ppFEV_1$ and NO_2 did not achieve statistically significant associations in the present study, which could be due to unmeasured variables or a lack of smoking data (Kelly & Fussell, 2011; McGarry *et al.*, 2019). To further understand how NO_2 affects lung function in pwCF, future research should take into account the smoking data as well as additional potentially confounding factors.

The impact of NO_2 on lung function in the CF population may be due to the confounding effect of traffic-related air pollution, which has a high degree of spatial variability and individuals

have much greater variability in traffic-related pollution exposure than $PM_{2.5}$ air pollutants. The effect of NO_2 on CF lung function may also vary with climatic factors such as air temperature and season (Szczesniak *et al.*, 2020). Warmer temperatures were associated with poor lung function in 1,378 participants enrolled in the U.S. CF Twin and Sibling Study (−0.34, 95%CI: -0.57; 0.10, p = 0.005) (Collaco *et al.*, 2011). The mechanisms by which temperature affects lung function are unknown, however, this could be due to infectious agents, aeroallergens, air pollution or socioeconomic status. Another possibility is that the interaction of temperature and air pollution may worsen lung function in people with CF. Also, lung function may vary by season, as patients could have higher lung function in the colder month of January compared to the warmer month of July. In the UK CF population, Qvist *et al.* found a higher lung function in spring compared to winter, however, the difference was not clinically significant (0.34%, 95%CI: 0.10; 0.59) (Qvist *et al.*, 2019). Thus, geographic and population-related factors might have impacted the magnitude of the effect of NO_2 on CF lung function.

In comparison to all pwCF, those who did not move (change their residential postcode) in the years 2012–2016 appeared to have stronger relationships between NO_2 quintiles and $ppFEV_1$ in children with CF. Stronger lower $ppFEV_1$ among non-movers compared to movers suggests that chronic NO_2 exposure may be more significant than recent exposure in children with CF. This also suggests that people with constant exposure are most at risk for the effects of air pollution. Although the exact mechanism by which NO_2 may harm the lungs and lung function is yet unknown, some explanations may include the fact that NO_2 is a free radical with the potential to create oxidative stress, which can harm cells and inflame the airways (Kelly, 2003). Also, NO_2 has been demonstrated to cause eosinophilic inflammation at high concentrations, increase epithelial damage, decrease mucin expression, and increase baseline smooth muscle tone in both animals and in vitro models (Ayyagari *et al.*, 2007).

There are large studies that looked at healthy individuals and did not find a relationship between lung function and NO_2. In a cross-sectional study conducted in Shanghai, China, Hou *et al.* did not find any statistical association between NO_2 and FEV_1 in 5,276 individuals aged ≥ 20 years (Hou *et al.*, 2020). The estimated effect on FEV_1 was 0.74 ml (95%CI: -17.13; 15.65) lower for every 10 $\mu g/m^3$ increase in NO_2 estimated by a validated satellite-based and land-use regression model (Xiao *et al.*, 2017). Similar results were obtained by another study, which

included 25,528 Swiss individuals. A 10 µg/m^3 increase in yearly mean NO_2 was associated with lower FEV_1, albeit this association was not statistically significant (-7.7 ml, 95%CI: 15:9; 0.5, p = 0.07) (Strassmann *et al.*, 2021). These findings could be the result of a different study design and exposure assignment. The negative effects of NO_2 on $ppFEV_1$ may have been confounded in pwCF by other clinical conditions such as smoking and chronic lung illnesses. Identifying the factors that modify susceptibility to the effects of air pollution in the CF population is important from the perspective of public health. More research is needed to fully explore the complex relationship between NO_2 and lung function in the CF population.

2.4.1.2 $ppFEV_1$ and $PM_{2.5}$

$PM_{2.5}$ has been associated with a range of adverse respiratory health effects, including changes in lung function (Li *et al.*, 2020). In the current study, it was found that children with CF had a lower $ppFEV_1$ in higher quintiles of $PM_{2.5}$ than in the lowest quintile (Q1); however, this was not a linear trend. The mean $ppFEV_1$ was 3.02% (95% CI: -5.15; - 0.93, p = 0.004) lower in the Q2 quintile of $PM_{2.5}$ than in the Q1 quintile in children with CF. There is no research on the relationship between $PM_{2.5}$ and lung function in children with CF. I could only identify one study that investigated how $PM_{2.5}$ affects the lung function of pwCF, but this included both children and adults with no stratification by age (Goss *et al.*, 2004). The two studies (Goss *et al.* and the current study) employed different methods for assigning air pollution exposure; the current study used a high-resolution surface air pollution model, while Goss *et al.,* used data from monitors that were placed close to the patients' residences.

My study result is consistent with Goss *et al.*'s study, where the authors reported a significant association between $ppFEV_1$ and $PM_{2.5}$ in the CF population. In their study of the US CF population, $ppFEV_1$ decreased by 0.5% (95%CI: 0.3; 0.9%) for every 10 µg/m^3 increase in $PM_{2.5}$ (Goss *et al.*, 2004).

The finding of an association between $PM_{2.5}$ and lung function is plausible, as $PM_{2.5}$ can penetrate deeply into the lung and irritate and damage the alveoli, which results in impaired lung function. The lungs are one of the main targets of $PM_{2.5}$-induced toxicity because they are the first places where $PM_{2.5}$ deposits in the airways, which causes airway inflammation, damages the bronchial mucociliary system, decreases bacterial clearance and impairs the immune response, leaving the lungs vulnerable to a variety of respiratory illnesses and

infections (Jia *et al.*, 2021). After being inhaled, $PM_{2.5}$ deposits in the pulmonary bronchi, alveoli, and surface of the airways, triggering oxidative stress, which is induced by activating the PI3K/Akt, NF-κb, Nrf2-KEAP1-ARE, JAK/STAT, and MAPK signalling pathways; and impairing the normal function of cells. It can even induce apoptosis and an inflammatory response, all of which result in respiratory damage and impaired lung function (Lu *et al.*, 2021; Pardo *et al.*, 2020; Xu *et al.*, 2020). According to animal research, prolonged exposures to $PM_{2.5}$ cause lung remodelling, oxidative stress, and inflammation (Hatzis *et al.*, 2006).

It has been extensively studied how air pollution affects healthy individuals and other respiratory conditions like COPD and asthma. Therefore, comparing the effects seen here in CF with those previously documented for other respiratory diseases can offer insight into the influence of air pollutants in CF and effective mitigation techniques. A recent meta-analysis that examined the relationship between particulate matter and lung function in healthy children aged 6 to 18 years included 25 studies from around the world, published up to May 2022, and found that FEV_1 decreased by 21.3 ml (95% CI: 13.9; 28.9) and 61.0 ml (95% CI: 25.8; 96.2) for every 10 $\mu g/m^3$ increase in short-term and long-term $PM_{2.5}$ exposure, respectively (Zhang *et al.*, 2022). Similarly, among healthy adults, a systematic review and meta-analysis of 33 studies published between 1964 and 2020 indicated that for every 10 $\mu g/m^3$ increase in short-term $PM_{2.5}$ exposure, FEV_1 decreased by 7.63 mL (95% CI: 10.62; 4.63) (da Silveira Fleck *et al.*, 2021). In a meta-analysis of asthmatic adults, Qibin *et al.* found a significant correlation between lower FEV_1/FVC ratios and $PM_{2.5}$ (0.20, 95%CI: 0.02;0.38, p = 0.03) (Qibin *et al.*, 2020). A large COPD population-based UK Biobank study found that for every 5 $\mu g/m^3$ increase in $PM_{2.5}$, FEV_1 decreased by 83.13 ml (95% CI: -92.50; -73.75) (Doiron *et al.*, 2019). All these results suggest a possible link between $PM_{2.5}$ and lung function, which is consistent with my findings of an association between $PM_{2.5}$ and $ppFEV_1$ in CF children. The possible reason for the lower lung function in children than in adults could be due to weak immune systems and underdeveloped lungs in children (Nicholas *et al.*, 2017). There was no evidence of a significant association between $ppFEV_1$ and $PM_{2.5}$ in adults with CF in the present study, but there is some unexpected trend in the opposite direction from what was expected. There was no evidence of a significant association between $ppFEV_1$ and $PM_{2.5}$ in adults with CF in the present study, but there is some unexpected trend in the opposite

direction from what was expected. That is difficult to explain, but this could be due to some unmeasured factors that were not accounted for in this study.

The evidence is not one way, as some studies did not find any statistically significant association between $PM_{2.5}$ and lung function (Gehring *et al.*, 2015; Takahashi *et al.*, 2006). For instance, in a longitudinal study conducted in Japan with 1,466 pre- and post-school children aged 8 to 12 years, Takebayashi *et al.* observed no statistical association between elevated $PM_{2.5}$ levels and lung function growth indicators (FEV_1, FVC and V50) (Takebayashi *et al.*, 2022). Gehring *et al.* reported no significant association between growth in FEV_1 and $PM_{2.5}$ concentration at birth or during follow-up among 555 participants at the ages of 8 and 12 years in the PIAMA (Prevention and Incidence of Asthma and Mite Allergy) birth cohort study in the Netherlands (Gehring *et al.*, 2015). These inconsistencies in the research findings may be due to variations in exposure assessment techniques, different socio-political and geographic conditions, sample sizes, diverse climates, varying compositions of ambient PM, weather conditions and participant characteristics. The association between exposure to pollution and lung function may be confounded by individuals who have poor lung function and more frequent exacerbations moving to urban areas to be nearer to specialised medical centres. Also, pwCF with more severe diseases may limit their exposure to outdoor air by spending most of their time indoors, which could confound the findings. To completely understand the association between $PM_{2.5}$ and lung function, especially in the adult CF population, more research is required.

2.4.2 Pulmonary exacerbations (PEx)

2.4.2.1 PEx and $PM_{2.5}$

Exacerbations of pulmonary disease are associated with a significant health burden both for patients and healthcare systems. As previously mentioned, there is no agreed standard definition of PEx in CF. PEx was defined differently in previous investigations, using the number of antibiotic courses needed for respiratory symptoms severe enough to require oral or intravenous antibiotic treatment (Goeminne *et al.*, 2013; Goss *et al.*, 2004; Jassal *et al.*, 2013). PEx was defined by Goeminne *et al.,* as two or more PEx requiring the administration of IV or oral antibiotics at home or in a hospital. Whereas according to Farhat *et al.,* a respiratory exacerbation was characterised as having three or more symptoms, including a

fever, increased sputum output or cough intensity, a change in sputum colour, worsening dyspnea, appetite loss, a > 10% drop in FEV_1, weight loss, or both. PEx was determined by Goss *et al.* to be two or more exacerbations requiring IV antibiotics at home or in a hospital. In contrast, I used the number of #IVdays to define PEx in the current study because I believe it is a more reliable and robust measure for observational studies, correlating closely to disease severity. Moreover, compared with oral antibiotic courses, which are poorly recorded in hospital records (as they may be prescribed outside of the hospital setting), the record of IV antibiotics tends to be much more reliable for entering the UK CF Registry.

The associations between PEx and $PM_{2.5}$ in children and adults with CF were different in the present study. In CF adults, the #IVdays increased by 1.11 (95%CI: 1.03; 1.21, p = 0.009) days in Q4 compared to Q1 of $PM_{2.5}$, demonstrating a statistically significant association between #IVdays and $PM_{2.5}$. Additionally, there were fewer CF adults (0.76, 95%CI: 0.62; 0.94, p = 0.010 in Q4) in the highest quintiles of $PM_{2.5}$ with zero IV days compared to Q1 of $PM_{2.5}$. In children with CF, I did not find the association that I was expecting. There was no evidence of an association between $PM_{2.5}$ and #IVdays in children with CF. I conducted a number of sensitivity analyses, such as a subgroup analysis of those whose residential postcodes remained the same between 2012 and 2016, as well as comparing individuals in the top and lowest 20^{th} percentiles of exposure to air pollutants. Additionally, I stratified the analysis by severe (homozygous *F508del*, *F508del*/MF, Gating/any (Except RF mutation), MF/MF) versus mild genotype (Residual (RF)/other, *F508del*/other, Other/Other, MF/other), but the results remained consistent. It is likely that this lack of association between $PM_{2.5}$ and #IVdays in children with CF could be due to unmeasured covariates in children. Children are more susceptible to the effects of air pollution because their lungs and immune defence mechanisms are still developing, and they may also be exposed to more air pollutants than adults due to higher minute ventilation, higher levels of physical activity, and more time spent outdoors (Kim, 2004). Therefore, it is important to investigate the relationship between air pollutants and disease severity in the high-risk CF population in future studies.

Globally, there is limited evidence to compare these results. Only two published studies so far have looked at the relationship between $PM_{2.5}$ and PEx in the CF population (Goss *et al.*, 2004; Jassal *et al.*, 2013). The first study (n= 11,484) enrolled pwCF over the age of six years who were registered in the CF Foundation National Patient Registry in the US in 1999 and 2000

(Goss *et al.*, 2004). The second study (n= 145) included CF children who were at least seven years old and had visited the CF centre at Children's Hospital in the US between 2004 and 2008 (Jassal *et al.*, 2013). In line with the present study's finding in CF adults, Goss and his colleagues have shown a significant association between $PM_{2.5}$ and PEx in the CF population. Though there was a difference in the unit measure of $PM_{2.5}$ in both studies, Goss *et al.*, used $PM_{2.5}$ as a continuous variable and I employed quintiles of $PM_{2.5}$ to determine whether there is a dose-response relationship between air pollutant exposure and PEx. They found that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with a 21% (95%CI: 7; 33%) increase in the odds of two or more exacerbations when adjusting for age, sex, weight, race, airway colonization with *P. aeruginosa* and *Burkholderia cepacia (B. cepacia)*, pancreatic function, and insurance status. Of the 145 patients included by Jassal *et al.*, 62% were between 6 and 14 years old and 93% were under 20 years old. Therefore, Jassal *et al.* study findings may be compared with the findings of our CF children's cohort. They did not find a statistically significant association between CF exacerbations and $PM_{2.5}$ (OR = 1.25, 95% CI: 0.59; 2.67, p = 0.555), which is consistent with the findings of the present investigation in children with CF. The current study's sample size (n = 8,965) is more comparable to that of the previous study by Goss *et al.* which was a high-powered study with 11,484. whereas the second study, which was conducted by Jassal *et al.* included a much smaller sample of 145 CF. These studies had a different age range; Goss *et al.* and Jassal *et al.* only included pwCF older than six years old, whereas I also included pwCF younger than six.

The finding of an association between $PM_{2.5}$ and #IVdays is plausible as pwCF' lungs have high levels of free myeloperoxidase and neutrophil-derived proteases and low levels of S-nitrosoglutathione which result in excessive oxidative stress and inflammation, and air pollutants may further irritate and harm the airways of pwCF (Berger, 2002; Snyder *et al.*, 2002). This could then impact the frequency and severity of airway infections in people with CF. In another study comparing the deposition and retention of technetium-99m-labelled iron oxide particles (5 microns) in pwCF and healthy volunteers, Brown *et al.* also found that pwCF had increased particulate deposition in the tracheobronchial areas of poorly ventilated lung regions (Brown *et al.*, 2001). These depositions may contribute to the explanation of the relationship between $PM_{2.5}$ and #IVdays.

The effects of air pollution on other respiratory diseases, such as COPD and asthma, have been well studied. Air pollution is associated with an increase in hospitalisation rates among COPD patients, and earlier research was conducted in different countries. For every 10 $\mu g/m^3$ increase in $PM_{2.5}$ levels, the probability of hospitalisation among COPD patients rose by 1.61% in the United States (Dominici *et al.*, 2006), 0.82% in Beijing (Gao *et al.*, 2019), 1.72% in Taiwan, and 6.87% in China (Tsai *et al.*, 2013). A systematic review and meta-analysis of 26 studies published in March 2021 in COPD patients found that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with a significantly higher risk of acute exacerbation (Effect Size = 1.01, 95%CI: 1.01;1.02, $p<0.001$) events (hospitalisation, emergency department visit, and outpatient visit) (Li *et al.*, 2022). Also, a meta-analysis that included 22 studies and 267,413 emergency room visits for hospital admission for asthma exacerbations supports the evidence that $PM_{2.5}$ is associated with asthma exacerbations (OR = 1.03; 95% CI: 1.01;1.05) (Orellano *et al.*, 2017). All these findings in individuals with COPD or asthma point to a potential relationship between $PM_{2.5}$ and lung function, so they might also be relevant for the CF population. Air pollutants can absorb chemical elements from the environment, which contribute to the formation of reactive oxygen species. The oxidative stress may induce damage to the airway epithelium, making the airways of people with respiratory diseases, including COPD and pwCF' more vulnerable to further damage and impairing the immune system's resistance capabilities. Another probable explanation is that air pollution might cause inflammation in the lungs, exacerbating COPD or pwCF already reduced pulmonary function.

In my opinion, one of the main strengths of my study is in relation to the precision of the exposure assignment. Goss *et al.* assigned yearly exposure by linking the centroid of the home zip code of pwCF to the air pollution readings from the nearest population-oriented monitors in the Aerometric Information Retrieval System database, which were within a 30-mile radius. Similarly, Jassal *et al.* used data from the closest monitoring station to assign the $PM_{2.5}$ exposure at the residential postal address. The present study was a lot more precise both in terms of the size of the residential areas considered (i.e., UK postcode *vs* US zip code) and in terms of exposure assignment. I used a newly produced high-resolution (25m X 25m) continuous modelled surface for annual average concentrations of $PM_{2.5}$ that matched the residential location of each individual.

There were also some differences in the variables adjusted in the regression model in these three studies. Goss *et al.* did not take genotype into account, but the other (Jassal *et al.*) used *CFTR* classes, which have limitations outlined in ***Section 1.2***. *CFTR* mutations are one of the important factors that correlate with the disease severity of CF. *CFTR* mutations, such as Class I-III considerably differ from class IV-V mutations in terms of pancreatic insufficiency, the number of exacerbations, and BMI, whereas class IV-V mutations are associated with a milder phenotype (Başaran *et al.*, 2021; Zielenski, 2000). Hence, we chose to adjust this variable in our final model. AIC was also improved after adding the genotype category to the regression model. In terms of nutrition, Jassal *et al.* did not account for height, weight, or BMI, whereas Goss *et al.* only corrected for weight in the final regression model. Since BMI is a widely recognised and used indicator for assessing the nutritional condition of pwCF and is strongly correlated with the severity of the CF disease, I adjusted my regression model estimates for BMI (Milla, 2007; Nagy *et al.*, 2022). One limitation that all three studies (two previous studies and the present study), had in common was that none of them adjusted for smoking behaviours in the regression model. All three studies have acknowledged the lack of inclusion of smoking data as a major limitation of their study. Smoking is particularly harmful to those with CF since it doubles the risk of lung injury and lung function decline in those who are already at risk of infections and tissue damage. In a cohort research study including 10,895 pwCF aged 6 to 18 who were enrolled from across US CF centres, the CF Foundation Patient Registry (CFFPR) reported a 2.4% drop in median ppFEV$_1$ with smoke exposure. Also, second-hand smoke was associated with reduced FEV$_1$ (9.8 percentile point decrease; $P < 0.001$) in pwCF (Collaco *et al.*, 2008; Oates *et al.*, 2020). Therefore, it is important to measure and take smoking into account in future CF population studies. In the UK, this would require the inclusion of more detailed and reliable data in the Registry.

2.4.2.2 PEx and NO_2

In my study, the level of annual NO_2 exposure had a significant spatial variation, ranging from 1.6 to 71.2 μg/m^3 among pwCF in GB in 2015. There was no evidence of a statistically significant association between NO_2 and PEx in children or adults with CF. To my knowledge, only three published studies have reported the relationship between NO_2 and PEx in the CF population (Farhat *et al.*, 2013; Goeminne *et al.*, 2013; Goss *et al.*, 2004). Goss *et al.* found no evidence of a significant relationship between NO_2 and PEx in pwCF living in the US (OR = 0.98,

95%CI: 091;1.07, p=0.719), and Farhat *et al.* found no evidence either in Brazil (RR = 1.26, 95%CI: 0.76; 2.09), which is consistent with the current study. Goeminne *et al.* however, reported an 11.6% ($p < 0.001$) increase in the likelihood of starting oral or IV for PEx in pwCF in Belgium. In comparison to Goss *et al.* and the current investigation, Goeminne *et al.* and Farhat *et al.* used smaller sample sizes of hundreds of individuals (215 and 103, respectively) but covered multiple years of data, whereas Goss *et al.* and my study used thousands (11,484 and 8,965 respectively). The age groups in these studies varied as well; Goss *et al.* included pwCF 6 years and older, while Farhat *et al.* and the current study included pwCF younger than 6 years. The age range of the participants in the Goeminne *et al.* study was not mentioned, however, the mean (SD) age was reported at 21(13) years. These studies used different methods to assign the data on air pollution exposure. While the current study used a high-resolution model surface air pollution map and Goeminne *et al.* used the kriging interpolation on 4X4 kilometre grids technique to assign the air pollution exposure, Goss *et al.* and Farhat *et al.* used annual summaries of air pollution data gathered from monitors located closest to patients' homes. It is plausible that NO_2 exposure increases the risk of acute lung tissue injury by causing inflammation of the bronchial mucosa and lowering the bronchoconstriction threshold (Ling & van Eeden, 2009).

A significant association between NO_2 and PEx has also been reported in studies on various other respiratory diseases. In a meta-analysis of 67 studies, 48 of which included data on children and 21 of which included data on adults, the relative risk for asthma exacerbations, defined as emergency room visits and hospital admissions, was found to be 1.014 (95% CI: 1.008; 1.020) for every 10 $\mu g/m^3$ increase in average 24-hour NO_2 (Zheng *et al.*, 2021). Similarly, Evangelopoulos *et al.* found that for every IQR rise in NO_2, the odds of an exacerbation increased by 16.4% (95% CI: 8.6; 24.6%) in COPD patients in the UK (Evangelopoulos *et al.*, 2021). These studies may suggest a link between NO_2 and PEx in CF.

The effect of air pollutants on PEx can vary depending on climate factors such as season, wind speed, pressure and relative humidity (Szczesniak *et al.*, 2020). Many COPD patients have experienced difficulty breathing in cold weather and high winds (Tseng *et al.*, 2013). A study in Germany reported a 1% increase in daily ambulatory visits with increasing wind speed (396.25 m/s) and a change in air pressure (0.85 K, 193.47 of log air pressure in Pa) among COPD patients (Ferrari *et al.*, 2012). In a COPD study in China, the relative risk of

hospitalisation increased by 1.07 (95%CI: 1.05; 1.09) for every 1% increase in relative humidity (Tian *et al.*, 2019). Flight *et al.* did not find a significant association between climate variables (temperature, humidity and vapour pressure) and the incidence of either PEx or antibiotic prescription among 98 adults with CF in Manchester (Flight *et al.*, 2014). This could be due to the small sample size. Future research with a large sample size is advised to validate the finding of the association between NO_2 and PEx in the CF population while accounting for all of these climate variables.

2.4.3 Strengths and limitations of my study

This is the first study of the health impact of air pollutants on pwCF in the UK. My study has several major strengths. First, the large number of pwCF studies provided good statistical power to assess the effect of NO_2 and $PM_{2.5}$ on disease severity in the CF population. Second, I used high-quality UK CF Registry data with regular annual review data, which is a large national database with excellent coverage and completion of detailed individual-level residential information (full residential postcode), to assess the impact of air pollution at the CF residential postcode area in Great Britain. Data quality is assured through a number of mechanisms, including mandatory data and range checks for clinically valid values, ensuring that only valid characters are entered. The UK CF Trust Registry team regularly conducts training sessions and site visits at CF centres to make sure the data entry team is aware of data entry guidelines and motivated to increase the completeness and accuracy of their data. Registry data managers and statisticians also perform a variety of data cleaning checks before performing statistical analysis. The patient's residential postcode is validated annually in the annual review visits by CF centres. Third, I used a higher-resolution (25m X 25m) continuous modelled surface for annual average concentrations of NO_2 and $PM_{2.5}$ in GB in 2015 data to estimate the $PM_{2.5}$ exposure for each pwCF's residential address. This method allowed me to overcome the space coverage limitations that typically occur when data is only obtained from monitoring stations. Additionally, using exposure data from the year before the yearly review date permitted an investigation of the influence of exposure on disease severity. Fourth, I used zonal analysis to assign the exposure at the postcode unit level, which is a lot more precise than what has been used in other studies of CF and environmental factors (Farhat *et al.*, 2013; Goeminne *et al.*, 2013; Goss *et al.*, 2004; Jassal *et al.*, 2013). Fifth, my stratified analysis based on age (children and adults) made it possible to observe the modified effect of

age on the relationship between NO_2 and $PM_{2.5}$ and disease severity. Finally, I have developed a novel improved genotype classification that more accurately categorises the severity of CF phenotypic manifestations and is also more compatible with available contemporary treatment approaches.

My study is not without limitations. First, Individual-level data on potential confounders, such as cigarette smoke exposure, was limited at the time of analysis; although some published research has found inconsistencies between smoking exposure and poor clinical outcomes in pwCF, none of these studies examined the impact of air pollution (Collaco *et al.*, 2008; Ortega-García *et al.*, 2012; Raju *et al.*, 2013). Second, information on pancreatic insufficiency was not available for the current project. However, residual mutation may have partially taken into account pancreatic status, which is strongly correlated with it and also has prognostic significance (Dos Anjos *et al.*, 2022; Poli *et al.*, 2020). Third, I did not factor in the diet in my regression analysis, and I doubt that this had any effect on the disease severity. I have adjusted my regression models for BMI, which is a reliable proxy measure for nutritional status and is frequently used to monitor and evaluate changes in an individual's nutritional status in clinics or hospitals. In children with CF, Zemel *et al.* found a strong correlation between dietary health and lung function (Zemel *et al.*, 2000). Fourth, my microbiological data has limitations as I only used a one-off sputum result, whereas using a definition for chronic pseudomonas infection may have been more relevant. However, I was unable to use the chronic infection definition due to the limited Registry data that was available for this project. Fifth, the cross-sectional nature of the study limits the ability to establish temporal causal links between air pollution exposure and lung function and PEx in the CF population. Moreover, without taking into account early exposures in the life course, I used modelled annual average concentrations of air pollution at individuals' home addresses at baseline as an individual-level long-term exposure to air pollution. I am aware of this limitation and to address this limitation, I did try to do a subgroup analysis among those who did not change their residential postcode between 2012 and 2016. In this CF population, 88.8% reported having two residential postcodes between 2007 and 2016, while only 60.0% had just one between those years. There was no significant change in the results among all pwCF and subgroups of pwCF who did not change their residential postcode between 2012 and 2016.

This was also reported by research from Scheers *et al.* who found that temporal variations in air pollution are much more determining than spatial differences (Scheers *et al.*, 2011).

Furthermore, for my CF population, I used the residential postcode as a proxy variable for the residence location. It is possible that pwCF might not live in this residential postcode. Thus, this could introduce a misclassification of the exposure; however, this misclassification would be more likely to be non-differential by disease severity.

Additionally, only outdoor exposure data for NO_2 and $PM_{2.5}$ was used for this project, making it unable to explore the cumulative effect of exposure to both outdoor and indoor air pollution. The outdoor air pollution may not correctly reflect how much air pollution pwCF are exposed to because the majority of pwCF may spend most of their time indoors. The indoor environment may influence disease severity (Lee *et al.*, 2020b), therefore, future studies should consider both outdoor and indoor air pollution for comprehensive analysis.

The small observed effect may reflect some unmeasured variables, such as host behaviour (the amount of time spent indoors and outdoors). Any registry study could be biased, particularly when it comes to unmeasured characteristics. Although major confounding variables such as age, sex, BMI, genotype, *P. aeruginosa*, deprivation, area of residence (urban vs rural) and country were considered in the analysis.

Although individual-level effect size on $ppFEV_1$ and #IVdays found in my research may seem small, population-level impacts can be significant. Even a slight decrease in lung function or increase in exacerbation days across a large population of pwCF can translate to substantial healthcare costs and burden. An increase of 1.11 (95% CI: 1.01-1.21, p=0.036) and 1.08 (95% CI: 0.89-1.17, p= 0.087) in #IVdays per year in the Q5 quintile of $PM_{2.5}$ before and after adjusting for various covariates implies a noteworthy relationship between elevated $PM_{2.5}$ concentrations and an increased number of intravenous (IV) days in CF adults. This study adds evidence to the potential public health significance of air pollution for CF management. However, the study also identifies limitations and uncertainties that warrant consideration in interpreting the results. First, Residual confounding remains a possibility. Despite adjusting for various covariates such as age, sex, BMI, lung function, *P. aeruginosa*, genotype, and deprivation, there may be unobserved confounding variables (such as smoking, and nutrition) that were not accounted for in the analysis. These unobserved factors could introduce bias

into the observed associations. Second, an ecological fallacy could be present. Attributing individual-level outcomes to aggregate exposure estimates might not accurately reflect CF's individual experiences. Third, the use of residential postcode exposure estimates obtained from land-use regression models introduces some variability in exposure assessment. This can potentially lead to exposure misclassification. Employing more accurate individual-level exposure assessments, such as personal monitoring devices, could enhance the precision of the effect estimates obtained in future studies.

Information on the patterns of daily activities of pwCF was not available. Also, neither household exposure nor exposure altering with time-related activities were taken into account. However, future research will need to use more advanced and sophisticated techniques to assess pwCF exposure more precisely and accurately. Other potential confounders that the regression analysis could not account for include comorbidities, family support, and physical exercise.

My study did not include distance to major roadways, as Jassal *et al.* found a significant association (OR=6.70, 95%CI: 1.23;54.49) between PEx (two or more exacerbations) and pwCF' residential proximity to major roadways in southern California (p = 0.042) (Jassal *et al.*, 2013). The possible explanation for the effect of this on CF exacerbations could be air pollutant toxicity to the pulmonary epithelium cells.

2.4.4 Conclusion

In conclusion, the results of my large GB CF study demonstrate that $PM_{2.5}$ is associated with a small increase in the risk of PEx in CF adults and a small reduction in lung function in CF children, while NO_2 was not shown to be associated with any of these outcomes in this CF population. My research contributes to the limited body of evidence of adverse associations between $PM_{2.5}$ and PEx in adults and lung function in children.

2.4.5 Recommendations

2.4.5.1 For future research

To further increase our understanding of the impact of NO_2 and $PM_{2.5}$ on disease severity in the CF population, future research on the complex relationship between exposure to

pollutants, including PM, but also microplastics and forever chemicals, and disease severity in susceptible CF populations should be conducted. Future research should make use of the Lung Clearance Index, one of the emerging physiologic biomarkers, to assess the condition of the airways in pwCF because it is a sensitive indicator of changes, especially in mild lung disease, a repeatable and useful measure of airway disease in CF, and more effective than conventional lung function tests at identifying early airway disease in children with CF (Horsley, 2009).

This research may help to increase our awareness and understanding of the harm that environmental pollutants cause to disease severity in the CF population and may provide a basis for establishing environmental protection laws that are specifically directed at vulnerable CF populations.

2.4.5.2 For pwCF

I recommend that pwCF be aware of the level of pollution where they live, both outdoors and indoors. These individuals may use air purifiers to reduce indoor exposure. A study found using an air purifier for three weeks significantly reduced indoor $PM_{2.5}$ concentrations (17 $\mu g/m^3$ *vs*. 9.26 $\mu g/m^3$, $p<0.001$), which resulted in reducing the frequency of medications used in children with asthma (Lee *et al.*, 2020a). Also, using real-time apps may help them avoid high-pollution routes and areas.

Use a face mask that is properly designed and closely fitted to block out pollution because using an inappropriate mask could give one a false sense of security and prolong exposure to high levels of ambient pollution (Ren *et al.*, 2016). Additionally, pay particular attention to their symptoms when smog levels are high, avoid areas with high pollution at specific times of the day and carefully plan the route when going outside.

3 Chapter III: Long-term (ten years) exposure to air pollutants and disease severity in the Greater London Area

3.1 Background

CF is a heterogeneous disease in which pulmonary disease is the main predictor of morbidity and mortality in the CF population (Liou *et al.*, 2001). Progressive decline in lung function is the hallmark of the CF clinical course (Szczesniak *et al.*, 2017a). A decline in lung function has been associated with increased morbidity and mortality in pwCF (Goss & Burns, 2007; Harun *et al.*, 2016).

In pwCF aged 18 to 51, Durda-Masny *et al.* found a significant association between life expectancy and $ppFEV_1$ ($p < 0.001$). When compared to pwCF with $ppFEV_1 > 70\%$, they found that those with $ppFEV_1 < 40\%$ had the lowest life expectancy (Durda-Masny *et al.*, 2021). Several factors including age, sex, height, weight, BMI (OR = 6.0 (95%CI: 5.0; 7.3), *CFTR* genotype, age at diagnosis, pancreatic function (OR = 2.0 (95%CI: 1.6; 2.5)), CF-related diabetes (CFRD) (OR = 1.84 (95%CI: 1.6; 2.2)), infection by *P. aeruginosa* sputum culture positive (OR = 2.4 (95%CI: 2.0; 2.7)) and number of PEx per year showed a statistically significant effect on $ppFEV_1$ in pwCF (all $p < 0.001$) (Eitan *et al.*, 2014). Numerous studies have shown that repeated PEx are associated with a steeper decline in lung function, specifically in FEV_1 in pwCF (Konstan *et al.*, 2007; Sanders *et al.*, 2011).

Although there has been substantial progress in understanding lung function trajectories in pwCF (Earnest *et al.*, 2020; Elpis *et al.*, 2023; Harun *et al.*, 2016; Macdougall *et al.*, 2023; Szczesniak *et al.*, 2017b; Taylor-Robinson *et al.*, 2012), there is still a research gap in understanding the impact of $PM_{2.5}$ on $ppEV_1$ trajectories in pwCF. Although its harmful effects on respiratory health are well known (Li *et al.*, 2018; Liu *et al.*, 2017; Sun *et al.*, 2020; Xing *et al.*, 2016; Zhang *et al.*, 2023), little research has been done on how $PM_{2.5}$ affects $ppFEV_1$ trajectories in the context of CF. To my knowledge, there has not been a longitudinal study performed to investigate the annual rate of change in $ppFEV_1$ associated with long-term exposure to $PM_{2.5}$ in pwCF over time (between 2008 and 2017). Understanding how $PM_{2.5}$ affects the $ppFEV_1$ trajectories in pwCF in London may help bridge this knowledge gap and

provide helpful insights for clinical practice and public health activities for pwCF. This study intends to fill the research gap and add to the body of knowledge by concentrating on the impact of $PM_{2.5}$ on $ppFEV_1$ trajectories in CF. The results of this study have the potential to understand the disease progression, guide targeted treatments and public health initiatives for pwCF in Greater London and beyond, with the ultimate goal of enhancing their respiratory health outcomes and general well-being.

3.1.1 Research question

Is there an association between $PM_{2.5}$ exposure and a decline in $ppFEV_1$ in the pwCF in Greater London?

3.1.2 Hypothesis

The decline in $ppFEV_1$ is stronger among pwCF who had high exposure compared to low exposure to $PM_{2.5}$.

3.1.3 Objective of the study

I aim to estimate the effect of $PM_{2.5}$ exposure on changes in lung function defined by studying longitudinal $ppFEV_1$ among a cohort of pwCF with a residential postcode in the Greater London Area (GLA) over a period of ten years (2008-2017).

3.2 Materials and methods

3.2.1 Study design

A longitudinal cohort study of pwCF with repeated measures was used to achieve the aim of this study.

3.2.2 Study geography and population

Using UK CF Registry data, all pwCF with a residential postcode in the GLA in 2017 were identified and selected. The GLA is located in southeast England and covers an area of 1,569 square kilometres (606 square miles) (London Online., 2023). The GLA includes the City of London as well as 32 London boroughs. All selected pwCF's ages were tracked back to 2008,

and those who were under the age of six in 2008 were excluded due to their inability to perform the standard pulmonary function tests accurately (Kerem *et al.*, 2014). Residential postcodes for the remaining pwCF for each year between 2008 and 2017 were taken from the UK CF Registry's postcode dataset. Although postcodes are meant to be recorded and updated at annual reviews, I identified some gaps and inconsistencies. I needed to have all postcodes between 2008 and 2017 for every pwCF in this study to determine the $PM_{2.5}$ exposure. To do this, I developed the following rules to impute the missing postcodes between 2008 and 2017 for all the pwCF.

i) A missing postcode was replaced with the following year's postcode if neighbouring postcodes were the same on both sides, for example, a postcode is missing for 2015, but the 2014 and 2016 postcodes were the same, so the missing postcode was replaced with the 2016 postcode.
ii) A missing postcode was replaced with the following year's postcode if neighbouring postcodes were different on both sides. See the below table for more details.
iii) Where two successive postcodes were missing, the first missing postcode was replaced with the previous year's postcode, and the second missing postcode was replaced with the following year's postcode.
iv) Where three postcodes in a row were missing, the first missing postcode was replaced with the previous year's postcode, the second missing postcode was replaced with the most common postcode during the study period, and the third missing postcode was replaced with the following year's postcode.
v) Where four postcodes in a row were missing, the first two missing postcodes were replaced with the previous year's postcode, 3rd and 4th missing postcodes were replaced with the following year's postcode.
vi) Where more than four postcodes were missing but not in a row, these were replaced using the above rule where appropriate.

Table 3.1: Explanation of replacing a missing postcode

ID	2008	2009	2010	2011	2012	2013	2014
X	a	a	Missing ->a	Missing -> b	b	b	b
Y	e	Missing -> e	Missing -> e	Missing -> b	Missing -> b	b	b
Z	h	h	Missing -> h	g	g	Missing -> c	c

ID column represents a patient
Capital letters in columns 2008 to 2014 represent a postcode.
Highlighted text is replaced with a postcode using the rule explained above.

3.2.3 Dummy postcode

To obtain specific exposure data for my pwCF with a residential postcode within Greater London, I had to share their postcode with Dr Benjamin Barratt, within the Environmental Research Group at the School of Public Health, Imperial College London. Following approval by the UK Cystic Fibrosis Trust and to preserve the confidentiality of the data, I generated a set of random dummy postcodes, a method commonly used for spatial data linkages. The aim of dummy data is to dilute real data with a given number of random replications so that it becomes virtually impossible to differentiate the real data from the dummy data. In this case, for every real pwCF's postcode, I generated ten random dummy postcodes. To make sure that these match the features of the real data, I looked at the frequency of postcode change throughout my study period (2008-2017). There were 40% of pwCF who did not change their postcode during the study period, and 32% changed their residential postcode once, as shown in ***Table 3.2*** below.

Table 3.2: Distribution of the number of unique postcodes between 2008 and 2017 (n=527)

Number of unique postcodes between 2008 and 2017	Number of pwCF	percentage of pwCF
01	209	39.7
02	170	32.3
03	80	15.2
04	35	6.6
05	19	3.6
06	10	-
07	<10	-

Then, I made sure that the distribution of changes in postcodes between 2008 and 2017 in my list of dummy postcodes matched the frequency shown above. The purpose of this

exercise was to ensure that my dummy postcode resembled real pwCF postcodes and that no one would be able to differentiate or identify a dummy pwCF postcode when this list was combined with a real pwCF postcode.

3.2.4 Outcome variable: Lung function ($ppFEV_1$)

FEV_1 measurements recorded in the UK CFR were standardised to $ppFEV_1$ based on age, sex, height and ethnicity using the GLI-2012 equations and NHANES III spirometry equations within the *rspiro* R package (Theodore Lytras., 2022). This package is a valuable and practical resource for those who need to calculate spirometry data. The *pctpred GLI* function from this r package was used to calculate the percent predicted FEV_1 for all the pwCF with a residential postcode in London between 2008 and 2017 using age, sex, height, ethnicity and FEV_1 data from the UK CF Registry. I used $ppFEV_1$ as a continuous variable to ensure a precise and accurate estimation of the result.

3.2.5 Exposure variable

The "London Air Pollution Toolkit" was used to model $PM_{2.5}$ exposure at the postcode level of pwCF with a residential postcode in London (Beevers *et al.*, 2012; Bos *et al.*, 2023). I selected this model because it is high-quality, has higher resolution, and has a 20 m x 20 m spatial resolution. It also models exposure at the postcode level, which is comparable to this study analysis as I was conducting it at the residential postcode level. This toolkit, a compressive instrument, was created to assess and calculate the level of air pollution in Greater London. Using a combined modelling-measurement approach and kernel modelling approach, this toolset predicts the dispersion of air pollution (Beevers *et al.*, 2012). This air pollution Toolkit's quality assurance and quality control procedures have previously been described in more depth (Kelly *et al.*, 2011; Whitehouse *et al.*, 2020). This model estimates the concentration of air pollution across London using detailed topography, meteorology, and emissions.

The London Air Pollution Toolkit was used to model the $PM_{2.5}$ exposure for all pwCF with a residential postcode in London for the following three periods.

i) yearly average exposure

ii) average exposure over a week preceding the annual review date

iii) and average exposure over a day preceding the annual review date

3.2.6 Covariates

I identified the following variables as potential confounders prior to the study based on their influence on lung function: age, sex, height, weight, *CFTR* genotype*, P. aeruginosa*, and deprivation (Townsend index) (Er *et al.*, 2020; Kerem *et al.*, 2014; Taylor-Robinson *et al.*, 2013). Each of these variables has been described in great depth above in ***Section 2.2.10.***

Using the new classification of mutation, I developed in this project; I have created a variable that differentiates between mild, Indeterminate and severe mutations. The mild mutation includes Residual (RF)/other, Indeterminate include *F508del*/other, Other/Other, and MF/other, whereas the severe mutation includes homozygous *F508del*, *F508del*/MF, Gating/any (Except RF mutations) and MF/MF.

3.2.7 Spatial and Statistical analysis

All statistical analysis was performed using *R* 4.2.2 and on SAHSU secure research systems (SSRS) (R Core Team., 2022). SSRS is a secure service platform that facilitates the handling and processing of sensitive personal data for research purposes. The Imperial College Secure Network uses practices that are ISO27001-approved to protect data storage.

A descriptive analysis was performed to explore the distribution of the demographic and clinical characteristics of pwCF. The mean, standard deviation, and range were reported for $ppFEV_1$, $PM_{2.5}$, and frequencies/proportions for age, sex, genotype, number of *P. aeruginosa* positive cultures and deprivation (Townsend index).

I used Spearman's correlation test to investigate the association between a repeated measure of $ppFEV_1$ and $PM_{2.5}$ during the study period using the "*ggpairs"* function in the "*ggally*" R package. The *ggpairs* allow for the creation of great scatterplots, which include the Pearson correlation coefficient, a pairwise comparison of each pair.

Trajectories of $ppFEV_1$ and $PM_{2.5}$ were calculated using "*brolgar"* (Browse over longitudinal data graphically and analytically in R) R package (Tierney. et. al., 2022). This package makes it

easier to explore longitudinal data and provides tools to identify and summarise interesting individual patterns in longitudinal data. Spaghetti plots were generated to explore the individuals' trajectories of $ppFEV_1$ and $PM_{2.5}$ over time. Stratified analyses of trajectory plots were created by age, sex and genotype to explore their potential modifying effects.

Data from the Registry are based on annual reviews. $ppFEV_1$ of pwCF is recorded while they are in a steady state and postponed if they do not meet this requirement. Registry records should therefore reflect the annual health progression of pwCF, but I wanted to test the hypothesis that short-term air pollution events could potentially affect these annual records. As a result, alongside the annual residential modelled exposure for $PM_{2.5}$ and NO_2, we also obtained average exposure to $PM_{2.5}$ and NO_2 for the week and the day (24 hours) preceding the annual review appointment. I did not have enough time to run and interpret the analyses for NO_2, so this chapter focuses on $PM_{2.5}$.

To investigate the relationship between $PM_{2.5}$ exposure and $ppFEV_1$, I divided $PM_{2.5}$ concentration into quintiles. However, due to the small numbers of pwCF in the lower (Q1) and top (Q5) quintiles of $PM_{2.5}$, I merged Q1 with Q2, and Q4 with Q5. So, I now have three exposure levels for my $PM_{2.5}$. low, medium, and high. Low exposure was defined as pwCF who were exposed to either Q1 or Q2 of $PM_{2.5}$ at least five times between 2008 and 2017. Similar to this, high exposure was applied to pwCF in London between 2008 and 2017, who were exposed to either Q4 or Q5 of $PM_{2.5}$ at least five times. All other pwCF were described as having a medium exposure if they were not classified as having a low or high $PM_{2.5}$ exposure.

Longitudinal analyses were employed to determine the annual rate of change in $ppFEV_1$ in pwCF associated with $PM_{2.5}$ exposure over time (2008 to 2017). Using the "*lmer"* function of the "*lme4*" R package, linear mixed models with random pwCF intercepts were used to estimate the decline in the $ppFEV_1$ over time associated with $PM_{2.5}$ exposure by including an interaction term between $PM_{2.5}$ quintile and time (follow-up year of the annual review) (Bates *et al.*, 2015). The model included a random intercept for each CF patient to account for CF individual differences and a dummy interaction term between $PM_{2.5}$ exposure level and year (follow-up year) to capture the differential change in $ppFEV_1$ across $PM_{2.5}$ exposure levels over time. The specific model employed is denoted as follows:

Eq1: $ppFEV_1 = \beta_1 \times yr_qun + \beta_2 \times age + \beta_3 \times sex + \beta_4 \times genotype + \beta_5 \times deprivation + \beta_6 \times height + \beta_7 \times weight + \beta_8 \times P.aeruginosa + (1 \mid id) + \varepsilon$

where:

- *ppFEV*$_1$ represents the predicted forced expiratory volume in one second.
- yr_qun represents the interaction between $PM_{2.5}$ exposure level (represented by quintiles) and year (follow-up year). It allows us to assess whether the annual rate of change in $ppFEV_1$ differs significantly across $PM_{2.5}$ exposure groups compared to a reference category
- $\beta_{2,}$ $\beta_{3,,}$ $\beta_{4,}$ $\beta_{5,}$ $\beta_{6,}$ and β_7 are the regression coefficients associated with each predictor of age, sex, genotype, deprivation index, height, weight, and presence of Pseudomonas aeruginosa infection, respectively.
- (1 | id) represents the random intercept term for each CF patient identified by a unique ID. This accounts for individual-specific variations in baseline lung function and any unobserved factors that may influence $ppFEV_1$ over time."
- ε represents the error term. It captures the random error or unexplained variability in the model

The model was adjusted for a set of a priori fixed and time-varying confounders based on biological knowledge. Fixed variables include sex, genotype and deprivation. Time-varying variables include age, height, weight and *P. aeruginosa.*

I chose to use the AIC criterion to identify the model with the best fit, i.e., the model that explained the most variation in $ppFEV_1$ with the fewest number of independent variables in the model. The final linear mixed model's underlying assumptions were checked using Q-Q plots of the residuals. The coefficient estimates are reported as a percentage change in $ppFEV_1$ with 95%CI for quintiles of $PM_{2.5}$ exposure. All results were considered statistically significant when the p-value was less than 0.05.

I performed various sensitivity analyses to confirm that the results were consistent and robust. First, I performed sensitivity analyses on the subset of pwCF who residential postcode did not change during the study period of 10 years (2008 – 2017). Second, stratified analyses

were performed by age group and genotype to explore their potential modifying effects on findings.

3.3 Results

Out of the 8,965 pwCF who had annual review in 2016, 527 pwCF had a residential postcode in London. Of these, a total of 134 were excluded from this analysis for the following reasons: 114 were ≤ 14 years in 2017 so these were under 6 years old in 2008, 15 had a lung transplant and 5 had missing exposure data (***Figure 3.1***). As a result, 393 pwCF were eligible for this trajectory analysis.

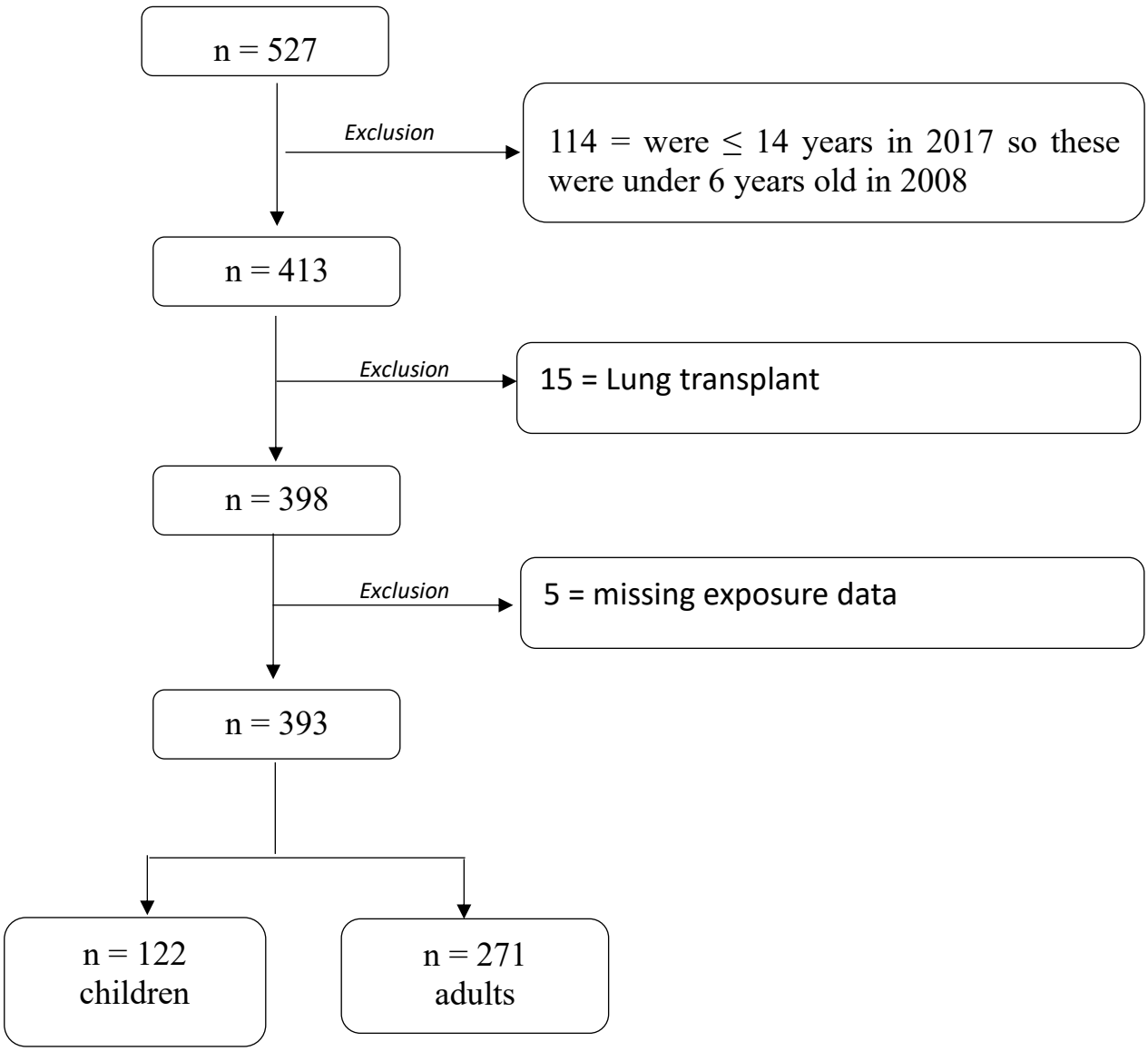


Figure 3.1: Flow diagram of the inclusion and exclusion criteria used for people with CF (pwCF) in the London UK CF Registry

3.3.1 Characteristics of pwCF for this analysis

Table 3.1 compares the demographic and clinical characteristics of pwCF with a residential postcode of Greater London and all pwCF in the UK CF registry in 2016. In this trajectory analysis cohort, there were 122 (30.4%) children between the ages of 6 and 15 years in 2008 and 271 (69%) were adults over the age of 16, which is a similar and comparable percentage to all pwCF in the UK CF registry in 2016 (p=0.775). There were 203 (51.7%) females in London which is a significantly higher percentage than all pwCF in the UK CF registry in 2016 (p=0.048). Similarly, almost half of pwCF (48.1%) was homozygous *F508del* in the London cohort compared to 49.7% in all pwCF in the UK CF registry in 2016. There were significant differences in the genotype categories under the revised genotyping classification, including F508del/MF (16% in the London cohort *vs*. 20% in the UK CF Registry cohort) and RF mutation (8% *vs*. 11% in the London and UK CF populations, respectively) ($p < 0.001$). There was no statistically significant difference in the genotype classification using the simplified classification of genotype ($p = 0.200$). Overall, the proportion of *P. aeruginosa* culture-positive in the London cohort increased from 29% in 2008 to 62% in 2017. In the UK CF registry cohort, there is likewise an upward trend from 28% in 2008 to 69%% in 2015 before it started to decline to 45% in 2017. With the exception of the years 2016 and 2017, when the London cohort's proportion of positive *P. aeruginosa* cultures was significantly higher ($p <0.001$), both cohorts' proportions of positive *P. aeruginosa* cultures were essentially the same. Townsend-based deprivation quintiles varied significantly between the two cohorts ($p <0.001$); in the London cohort, there were 6% (n=23) pwCF who lived in the least deprived Townsend index area and 43.5% (171) in the most deprived area. However, there was an equal distribution of pwCF across all Townsend index quintiles in the UK CF registry cohort.

Table 3.3: Comparison of characteristics of pwCF with a residential postcode in the Greater London Area (GLA) and all pwCF in the UK CF Registry in 2016

	pwCF in GLA n=393	**All CF in 2016 in the UK CF registry n=8,965**	**p-value**
Characteristics	n (%)	**n (%)**	
Age (years)			0.775
Children (6 – 15)	122 (31.0)	2,268 (30.4)	
Adults (≥ 16)	271 (69.0)	5,203 (69.6)	
Sex			**0.048**
Female	203 (51.7)	4,175 (46.7)	
Male	190 (48.3)	4,790 (53.3)	
Genotype			**<0.001**
homozygous F508del	189 (48.1)	4,456 (49.7)	
F508del/MF	62 (15.8)	1,757 (19.6)	
RF/other	32 (8.1)	980 (10.9)	
F508del/other	34 (8.7)	631 (7.0)	
Gating/any**	28 (7.1)	555 (6.2)	
MF/MF	23 (5.9)	291 (3.3)	
Other/other	13 (3.3)	157 (1.8)	
MF/other	12 (3.1)	138 (1.5)	
Genotype			**0.200**
homozygous F508del	189 (48.1)	4,456 (49.7)	
heterozygous F508del	96 (24.4)	2,388 (26.6)	
Other	108 (27.5)	2,121 (23.7)	
P. aeruginosa* culture *positive			
2008	113 (28.8)	2,055 (27.7)	0.637
2009	152 (38.7)	2,452 (37.6)	0.661
2010	198 (50.4)	2,927 (45.4)	0.052
2011	205 (52.2)	3,347 (54.7)	0.330
2012	213 (54.2)	3,820 (58.9)	0.064
2013	213 (54.2)	3,390 (62.4)	**0.001**
2014	225 (57.3)	4,252 (67.8)	**<0.001**
2015	230 (58.5)	4224 (69.3)	**<0.001**
2016	240 (61.1)	4,139 (44.8)	**<0.001**
2017	242 (61.6)	4,352 (44.5)	**<0.001**
Deprivation quintile (Townsend index)			**<0.001**
Q1 (least deprived)	23 (5.9)	1,855 (20.7)	
Q2	34 (8.7)	1,875 (20.9)	
Q3	85 (21.6)	1,863 (20.8)	
Q4	80 (20.4)	1,784 (19.0)	

Q5 (most deprived)	171 (43.5)	1,588 (17.7)

pwCF= people with cystic fibrosis; MF = Minimal function (a mutation that produces either no protein or protein that does not respond to currently ved *CFTR* modulators); RF = Residual function; Other = Either the mutation was not identified, or it was identified but classification into one of the categories was not possible. **Except RF mutation

3.3.2 Summary statistics of $ppFEV_1$ and $PM_{2.5}$

Table 3.2 shows the summary statistics of mean $ppFFEV_1$ trajectories in pwCF with a residential postcode in Greater London between 2008 and 2017. The mean $ppFEV_1$ ranged from 63.1% to 73.3% with a standard deviation of 22.6 to 26.2%. Similarly, the median range of the $ppFEV_1$ was between 63.6 and 75.5%, suggesting that the mean $ppFEV_1$ distribution was symmetrical. During the study period, $ppFEV_1$ values ranged between 12.9% and 143%. On average, the mean $ppFEV_1$ trajectories decreased by 10% during the ten years from 2008 to 2017, or about a 1% decrease annually ***Figure 3.2***.

Table 3.4: Summary of $ppFEV_1$ measurements in pwCF with a residential postcode in the GLA between 2008 and 2017

Year	N	Mean	SD	Median	Q1	Q3	Min	Max
2008	301	73.3	22.6	75.5	58.5	87.6	20.6	129
2009	338	72.6	22.7	74.1	57.3	86.1	18.7	134
2010	351	70.9	23.0	71.5	54.7	86.4	19.2	131
2011	373	71.1	24.2	72.6	54.5	87.8	17.1	143
2012	381	69.0	23.5	69.7	50.5	86.2	15.2	126
2013	379	67.6	23.1	67.4	49.6	84.7	19.8	128
2014	384	67.6	26.2	65.9	46.8	86.0	12.9	135
2015	371	65.8	22.7	67.0	48.2	84.1	14.9	134
2016	376	65.0	24.4	64.8	47.3	82.8	13.4	127
2017	384	63.1	25.2	63.6	43.1	80.8	13.0	137

$ppFEV_1$ = percent predicted FEV_1; pwCF = people with cystic fibrosis; SD= standard deviation.



Figure 3.2: Overall mean ppFEV$_1$ trend among pwCF with a residential postcode in the GLA between 2008 and 2017

Table 3.3 and ***Figure 3.3*** summarise the mean $PM_{2.5}$ concentrations of the yearly, weekly and daily before the annual ppFEV$_1$ measurement in pwCF living in Greater London between 2008 and 2017. The annual mean $PM_{2.5}$ concentrations varied between 13.6 and 14.0 µg/m^3 between 2008 and 2017, while the mean $PM_{2.5}$ concentrations in weekly readings ranged from 12.9 to 14.1 µg/m^3. Additionally, the average $PM_{2.5}$ concentrations daily varied between 12.4 and 14.2 µg/m^3 during the study period. The range of annual mean concentration of $PM_{2.5}$ varied from 11.5 to 17.3 µg/m^3 during the study period. This range varied more during weekly and daily than before the annual ppFEV$_1$ measurement. A week prior, $PM_{2.5}$ concentrations ranged from 5.4 to 45.5 µg/m^3, and a day prior it was between 3.3 and 65.9 µg/m^3. Over the course of the research period, the mean (SD) $PM_{2.5}$ concentrations were 13.3 (8.2), 13.5 (5.7) and 13.9 (1.1) µg/m^3 on a daily, weekly, and yearly, respectively. My analyses showed that the level of average exposure for the three time periods, daily, weekly and annually, were very consistent for our cohort of pwCF (***Figure 3.3***). I, therefore, opted to present results only for the annual period as this was the most

consistent with my original hypothesis and the yearly frequency of annual reports in the Registry.

Table 3.5: Summary of yearly, weekly and daily pre-annual review estimated concentration of $PM_{2.5}$ at the residential postcode of pwCF in the GLA between 2008 and 2017

Year	N	Mean	SD	Median	Q1	Q3	Min	Max
$PM_{2.5}$ yearly average concentration before the annual $ppFEV_1$ measurement								
2008	364	14.0	1.1	13.8	13.1	14.8	11.7	16.4
2009	327	13.8	1.1	13.7	13.0	14.7	11.6	16.7
2010	342	13.8	1.2	13.9	12.9	14.7	11.6	16.8
2011	335	14.0	1.2	13.9	13.1	14.9	11.8	16.7
2012	334	14.0	1.2	14.0	13.1	14.9	11.5	16.8
2013	327	14.0	1.1	13.8	13.0	14.8	11.8	16.9
2014	350	13.9	1.1	13.8	13.0	14.8	12.0	17.3
2015	346	13.7	1.1	13.4	12.8	14.6	11.7	16.6
2016	339	13.6	1.1	13.4	12.7	14.6	11.4	15.9
2017	353	13.8	1.1	13.6	12.9	14.6	11.9	16.1
$PM_{2.5}$ weekly average concentration before the annual $ppFEV_1$ measurement								
2008	363	14.1	6.2	12.4	9.8	16.6	5.6	41.1
2009	324	13.9	5.9	12.2	9.8	16.6	6.4	35.8
2010	342	13.3	6.1	11.4	9.1	15.9	5.7	45.5
2011	334	13.5	5.7	11.7	9.5	15.6	6.4	41.8
2012	333	13.5	5.4	11.6	9.7	15.6	5.5	35.2
2013	326	12.9	5.4	11.2	9.3	14.5	5.4	41.9
2014	350	13.1	5.4	11.6	9.6	15.2	5.4	41.7
2015	345	13.5	5.6	11.9	9.5	15.9	5.7	33.6
2016	339	13.3	6.0	11.7	9.5	15.4	5.8	43.5
2017	353	13.6	5.7	12.0	9.8	15.6	6.2	37.3
$PM_{2.5}$ daily average concentration before the annual $ppFEV_1$ measurement								
2008	364	14.1	8.7	11.4	8.7	16.3	4.0	65.9
2009	327	14.2	9.5	11.1	8.3	16.7	3.3	60.8
2010	342	13.3	8.8	10.8	8.1	15.5	4.1	60.0
2011	335	13.1	8.0	11.0	8.3	14.8	3.5	57.1
2012	334	12.8	7.3	10.9	8.3	14.9	4.5	63.9
2013	327	12.7	7.2	10.2	8.0	14.3	3.3	48.4
2014	350	12.4	6.6	10.3	8.2	13.9	3.8	48.6
2015	346	13.6	8.2	10.9	8.2	16.3	3.5	47.2
2016	339	13.1	8.3	10.6	8.0	15.3	3.5	64.7
2017	353	13.4	8.6	10.6	8.2	15.3	3.3	61.2

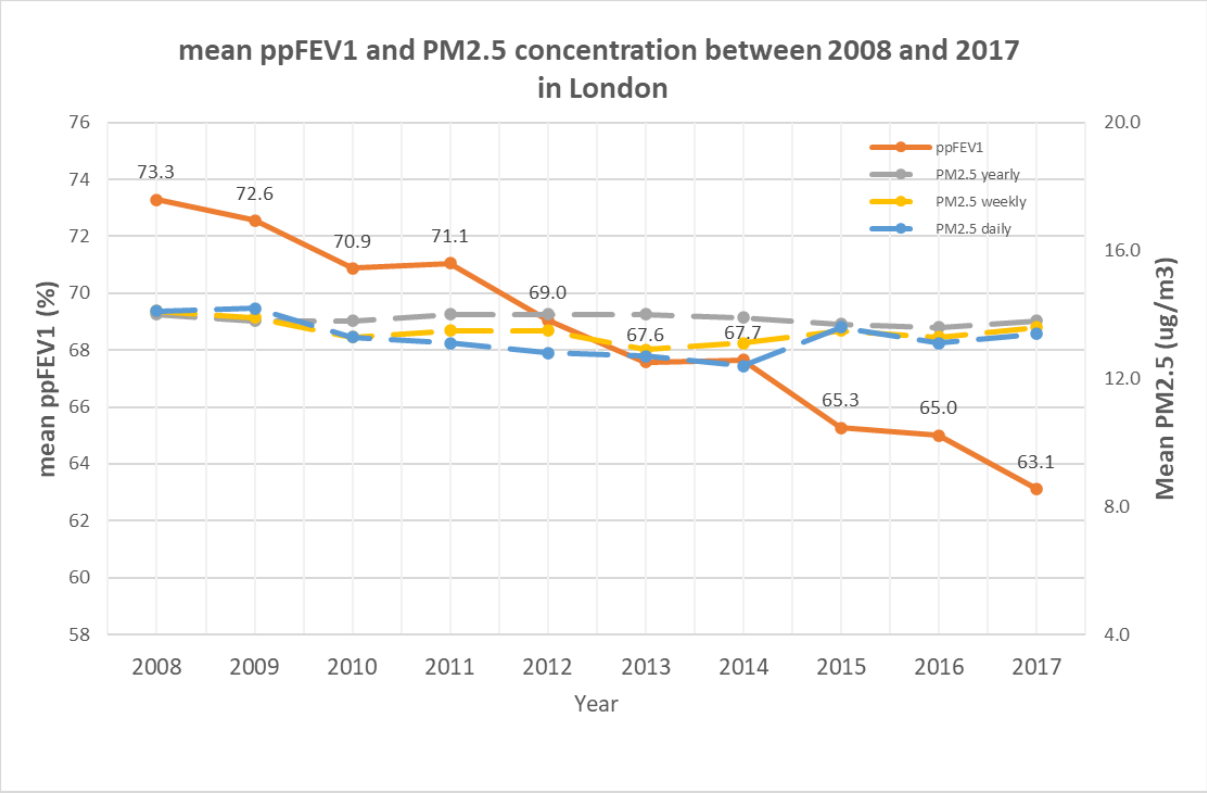


Figure 3.3: Mean $ppFEV_1$ and yearly, weekly and daily $PM_{2.5}$ concentration between 2008 and 2017 in the GLA

3.3.3 Correlation within $ppFEV_1$ and $PM_{2.5}$

The repeated measure of mean $ppFEV_1$ measurements showed strong correlations between the mean $ppFEV_1$ measurements between 2008 and 2017, as shown in the correlation plot ***Figure 3A*** (see ***Appendix 5.2***). The Spearman correlation coefficients ranged from 0.79 to 0.95. A similar correlation pattern was observed when yearly estimates of $PM_{2.5}$ exposure between 2008 and 2017 were explored, as shown in ***Figure 3B Appendix 5.2***. The Spearman correlation coefficients for $PM_{2.5}$ exposure ranged from 0.77 to 0.98, indicating a significant association between the measured yearly $PM_{2.5}$ exposures throughout the study period.

3.3.4 Longitudinal analysis of $PM_{2.5}$ and $ppFEV_1$

Table 3.4 shows the primary analysis of unadjusted and adjusted estimated change in mean $ppFEV_1$ trajectories associated with annual exposure of $PM_{2.5}$ among pwCF with a residential postcode within Greater London between 2008 and 2017. The results of the linear mixed models show a significant linear decline in mean $ppFEV_1$ during the subsequent

follow-up years; this decline is more pronounced among pwCF with high exposure to $PM_{2.5}$ than the low exposure group between 2008 and 2017.

In 2008 at the baseline of this study, pwCF with high exposure to $PM_{2.5}$ had a mean $ppFEV_1$ that was 3% lower than that of the group with low exposure. This decline was 9% in 2012 and 16% in 2017 among pwCF with high exposure. In my unadjusted linear mixed model, the mean decline in $ppFEV_1$ was -16.3% (95%CI: -24.40; -8.20) in higher exposed compared to those with low exposed which was -7.50% (95%CI: -10.70; -4.34) in 2017. This decline was consistent and remained significant after adjusting for age, sex, genotype, height, weight, *P. aeruginosa* and deprivation -14.5%, (95%CI: -22.0, -7.04) vs -7.56% (95%CI: -11.00; -4.10). ***(Figure 3.4)*.** The AIC values are summarised in ***Table 3.7*** along with the effect of the covariate on the regression model's goodness of fit. While including all potential confounders in a model might seem ideal for capturing all influences, it can lead to overfitting and multicollinearity issues. These issues can inflate standard errors and make it difficult to obtain reliable estimates of effect sizes. To address this, a two-step approach was employed for model selection. First, a comprehensive "full model" was built that included all potential confounders. Subsequently, the exposure variable ($PM_{2.5}$) was added to this full model. A log-likelihood ratio test was then conducted to assess the improvement in model fit after including the exposure variable. The resulting statistically significant p-value (<0.001) indicated that the model with $PM_{2.5}$ provided a significantly better fit than the model without it. This suggested that $PM_{2.5}$ exposure had a substantial and independent association with $ppFEV_1$, even after accounting for other contributing factors.

Table 3.6: Estimated change in mean $ppFEV_1$ associated with annual exposure of $PM_{2.5}$ among pwCF with a residential postcode in the GLA between 2008 and 2017 (n= 393, $ppFEV_1$ measurements = 3,333)

	Crude			Adjusted[2]		
Characteristic	**Beta**	**95% CI[1]**	**p-value**	**Beta**	**95% CI[1]**	**p-value**
Interaction between year and $PM_{2.5}$ concentration (μg /m^3)						
2008 * low exposure	Ref	—		Ref	—	
2008 * middle exposure	0.40	-5.93, 6.72	0.902	1.65	-4.22, 7.53	0.581
2008 * high exposure	-2.98	-11.2, 5.21	0.475	-0.69	-8.27, 6.88	0.857
2009 * low exposure	-2.72	-5.96, 0.52	0.099	-2.44	-5.87, 0.99	0.163
2009 * middle exposure	-0.73	-7.04, 5.58	0.820	0.55	-5.31, 6.41	0.854
2009 * high exposure	-5.23	-13.4, 2.89	0.206	-3.00	-10.5, 4.51	0.433
2010 * low exposure	-2.67	-5.91, 0.58	0.108	-2.47	-6.03, 1.09	0.175
2010 * middle exposure	-1.95	-8.25, 4.35	0.543	-0.21	-6.08, 5.66	0.943
2010 * high exposure	-7.34	-15.5, 0.80	0.077	-5.20	-12.7, 2.33	0.176
2011 * low exposure	-3.97	-7.14, -0.80	**0.014**	-4.74	-8.12, -1.35	**0.006**
2011 * middle exposure	-2.18	-8.47, 4.12	0.498	-0.66	-6.51, 5.19	0.825
2011 * high exposure	-7.61	-15.7, 0.50	0.066	-5.29	-12.9, 2.28	0.170
2012 * low exposure	-2.56	-5.76, 0.63	0.116	-2.30	-5.71, 1.12	0.187
2012 * middle exposure	-4.71	-11.0, 1.58	0.142	-3.55	-9.41, 2.30	0.234
2012 * high exposure	-8.76	-16.9, -0.68	**0.034**	-6.72	-14.2, 0.78	0.079
2013 * low exposure	-4.10	-7.25, -0.94	**0.011**	-4.68	-8.06, -1.30	**0.007**
2013 * middle exposure	-5.72	-12.0, 0.57	0.075	-4.45	-10.3, 1.40	0.136
2013 * high exposure	-11.0	-19.1, -2.90	**0.008**	-9.23	-16.7, -1.71	**0.016**
2014 * low exposure	-4.56	-7.72, -1.39	**0.005**	-5.19	-8.64, -1.74	**0.003**
2014 * middle exposure	-5.61	-11.9, 0.68	0.080	-4.55	-10.4, 1.30	0.127
2014 * high exposure	-11.0	-19.1, -2.88	**0.008**	-9.14	-16.6, -1.64	**0.017**
2015 * low exposure	-7.80	-11.0, -4.63	**<0.001**	-7.31	-10.7, -3.95	**<0.001**
2015 * middle exposure	-7.82	-14.1, -1.52	**0.015**	-6.44	-12.3, -0.60	**0.031**
2015 * high exposure	-12.2	-20.3, -4.09	**0.003**	-10.4	-17.9, -2.91	**0.007**
2016 * low exposure	-6.12	-9.31, -2.93	**<0.001**	-6.04	-9.39, -2.70	**<0.001**
2016 * middle exposure	-8.65	-14.9, -2.36	**0.007**	-7.23	-13.1, -1.39	**0.015**
2016 * high exposure	-13.7	-21.8, -5.56	**<0.001**	-11.3	-18.8, -3.81	**0.003**
2017 * low exposure	-7.50	-10.7, -4.34	**<0.001**	-7.56	-11.0, -4.10	**<0.001**
2017 * middle exposure	-10.6	-16.9, -4.34	**<0.001**	-9.05	-14.9, -3.20	**0.002**
2017 * high exposure	-16.3	-24.4, -8.20	**<0.001**	-14.2	-21.7, -6.78	**<0.001**

$ppFEV_1$ = percent predicted with FEV_1,

Low exposure of $PM_{2.5}$= pwCF who were exposed to either Q1 or Q2 of $PM_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

High exposure of $PM_{2.5}$= pwCF who were exposed to either Q4 or Q5 of $PM_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

Medium exposure of $PM_{2.5}$= all others who were not classified in either low or high exposure of $PM_{2.5}$

[1]CI = Confidence Interval

[2]Adjusted for age at baseline, sex, genotype, height, weight, *P. aeruginosa* and deprivation (Townsend index)

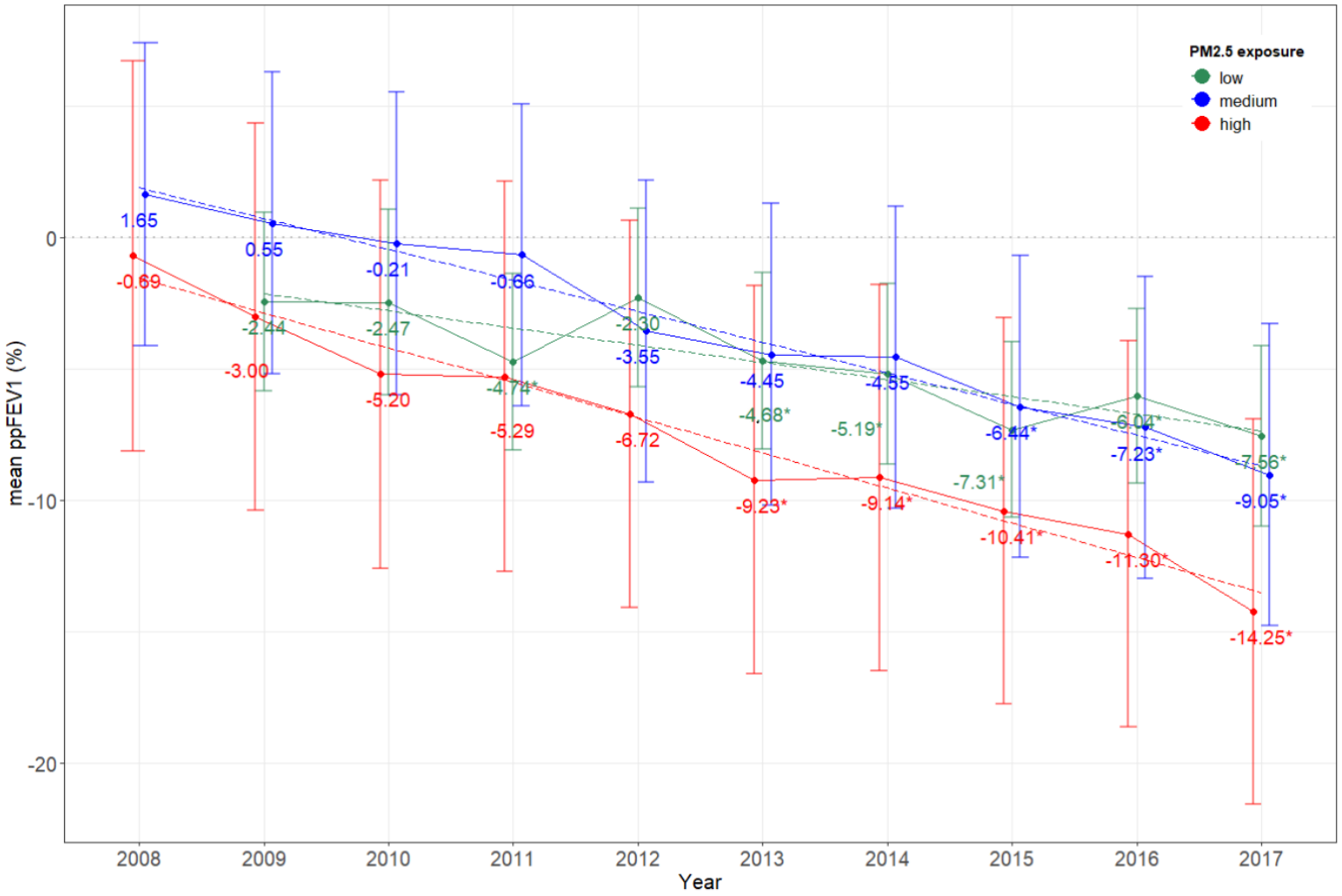


Figure 3.4: Comparison of the adjusted* estimated change in mean ppFEV$_1$ by exposure of PM$_{2.5}$ among pwCF with a residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, genotype, height, weight, *P. aeruginosa* and deprivation (Townsend index)

Table 3.7: Summary of the change in AIC with the addition of a covariate in the regression model

In the model	**AIC**
Interaction Year * $PM_{2.5}$ exposure	27868.6
Interaction Year * $PM_{2.5}$ exposure + Age	27841.5
Interaction Year * $PM_{2.5}$ exposure + Age + Sex	27788.5
Interaction Year * $PM_{2.5}$ exposure + Age + Sex + Genotype	27775.5
Interaction Year * $PM_{2.5}$ exposure + Age + Sex + Genotype + Height	25600.8
Interaction Year * $PM_{2.5}$ exposure + Age + Sex + Genotype +Height + Weight	25561
Interaction Year * $PM_{2.5}$ exposure + Age + Sex + Genotype +Height + Weight + *P. aeruginosa*	25561.5
Interaction Year * $PM_{2.5}$ exposure + Age + Sex + Genotype+ Height + Weight + *P. aeruginosa* and *Townsend index* (deprivation)	25566.1

3.3.5 Sensitivity /Stratified analysis

I performed various sensitivity and subgroup analyses to validate the robustness of my primary results. First, I performed a subgroup analysis of pwCF in London who did not change their postcode between 2008 and 2017. I did not find a clear linear trend in the reduction in $ppFEV_1$ among this subset of pwCF who were exposed to different levels of $PM_{2.5}$, as found in all pwCF analyses. The mean $ppFEV_1$ decreased by 15% (95% CI: -29.1; -0.88) in the high exposure group compared to -9.15% (95% CI: -14.40; -3.84) in the low exposure group in 2017, and the association was attenuated when the model was adjusted for possible confounders -9.3% (95% CI: -22.7; 4.12) vs -16.3% (95% CI: -16.3; -4.29) ***(Table 3.8 and Figure 3.5 & 3.6)***.

Second, I did a stratified analysis by children and adults and found that high $PM_{2.5}$ exposure had a stronger negative impact on $ppFEV_1$ in adults with CF. The decline was 13.0% (95% CI; -22.4, -3.65) in CF adults in the high-exposure group compared to -5.46% (95% CI; -9.15, -1.76) in the low-exposure group in 2017. There was no clear linear pattern in the reduction in $ppFEV_1$ among children with CF ***(Table 3.10 &3.11, Figure 3.9-3.10).***

Third, I employed the analysis in a subset of pwCF with a severe genotype (homozygous *F508del*, *F508del*/Minimal Function (*MF*), *Gating* and *MF/MF*). I found the mean $ppFEV_1$ was 12.8% lower (95% CI: 21.6; -3.99, p-value 0.004) in the high exposure group in 2017 compared to pwCF in the low exposure group in 2008, This association was consistent after adjusting for age, sex, height, weight, *P. aeruginosa* and deprivation (11.8%, 95%CI: -19.7; -3.86) vs (6.61%, 95%CI: -10.5; -2.67) ***(Table 3.11 & Figure 3.11 and 3.12)***.

Table 3.8: Estimated change in mean $ppFEV_1$ associated with annual exposure of $PM_{2.5}$ among pwCF who did not change their postcode in the GLA between 2008 and 2017 (n= 144, $ppFEV_1$ measurements = 1,225)

		Crude			Adjusted[2]	
Characteristic	**Beta**	**95% CI[1]**	**p-value**	**Beta**	**95% CI[1]**	**p-value**
Interaction between year and $PM_{2.5}$ concentration (μg /m^3)						
2008 * low exposure	Ref	—		Ref	—	
2008 * middle exposure	1.98	-8.71, 12.7	0.715	4.03	-6.31, 14.4	0.443
2008 * high exposure	-3.04	-17.2, 11.2	0.673	2.39	-11.1, 15.9	0.727
2009 * low exposure	-3.28	-8.76, 2.21	0.241	-2.92	-8.84, 3.00	0.333
2009 * middle exposure	2.21	-8.46, 12.9	0.684	4.13	-6.17, 14.4	0.430
2009 * high exposure	-6.78	-21.0, 7.46	0.349	-1.41	-15.0, 12.1	0.837
2010 * low exposure	-2.57	-7.99, 2.85	0.352	-3.35	-9.34, 2.65	0.273
2010 * middle exposure	0.48	-10.2, 11.1	0.930	2.57	-7.76, 12.9	0.624
2010 * high exposure	-10.7	-24.9, 3.55	0.140	-5.45	-19.1, 8.15	0.430
2011 * low exposure	-3.84	-9.23, 1.56	0.163	-5.09	-10.9, 0.75	0.087
2011 * middle exposure	0.48	-10.2, 11.1	0.929	3.16	-7.14, 13.5	0.546
2011 * high exposure	-10.3	-24.5, 3.86	0.153	-6.11	-19.7, 7.50	0.377
2012 * low exposure	-1.76	-7.21, 3.69	0.527	-1.93	-7.94, 4.09	0.530
2012 * middle exposure	-2.48	-13.1, 8.18	0.647	-0.33	-10.6, 9.98	0.950
2012 * high exposure	-11.2	-25.3, 2.91	0.119	-5.36	-18.8, 8.07	0.432
2013 * low exposure	-3.98	-9.28, 1.32	0.141	-5.58	-11.6, 0.45	0.070
2013 * middle exposure	-3.35	-14.0, 7.31	0.536	-1.45	-11.8, 8.84	0.781
2013 * high exposure	-10.9	-25.0, 3.17	0.128	-4.89	-18.4, 8.57	0.474
2014 * low exposure	-4.69	-10.0, 0.66	0.086	-6.83	-13.3, -0.37	**0.038**
2014 * middle exposure	-4.25	-14.9, 6.41	0.432	-2.15	-12.5, 8.16	0.682
2014 * high exposure	-10.1	-24.2, 4.01	0.160	-4.75	-18.2, 8.71	0.487
2015 * low exposure	-7.66	-13.0, -2.36	**0.005**	-7.07	-13.0, -1.16	**0.019**
2015 * middle exposure	-6.74	-17.4, 3.92	0.214	-4.61	-14.9, 5.68	0.378
2015 * high exposure	-13.3	-27.5, 0.94	0.067	-8.74	-22.3, 4.81	0.205
2016 * low exposure	-4.78	-10.2, 0.61	0.082	-5.54	-11.3, 0.28	0.062
2016 * middle exposure	-7.27	-17.9, 3.38	0.180	-5.31	-15.6, 4.97	0.310
2016 * high exposure	-12.6	-26.7, 1.55	0.081	-7.05	-20.4, 6.34	0.300
2017 * low exposure	-9.15	-14.4, -3.84	**<0.001**	-10.3	-16.3, -4.29	**<0.001**
2017 * middle exposure	-9.01	-19.7, 1.64	0.097	-6.43	-16.7, 3.86	0.219
2017 * high exposure	-15.0	-29.1, -0.88	**0.037**	-9.31	-22.7, 4.12	0.173

$ppFEV_1$ = percent predicted with FEV_1,

Low exposure of $PM_{2.5}$= pwCF who were exposed to either Q1 or Q2 of $PM_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

High exposure of $PM_{2.5}$= pwCF who were exposed to either Q4 or Q5 of $PM_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

Medium exposure of $PM_{2.5}$= all others who were not classified in either low or high exposure categories

[1]CI = Confidence Interval

[2]Adjusted for age at baseline, sex, height, weight, genotype, *P. aeruginosa* and deprivation (Townsend index)

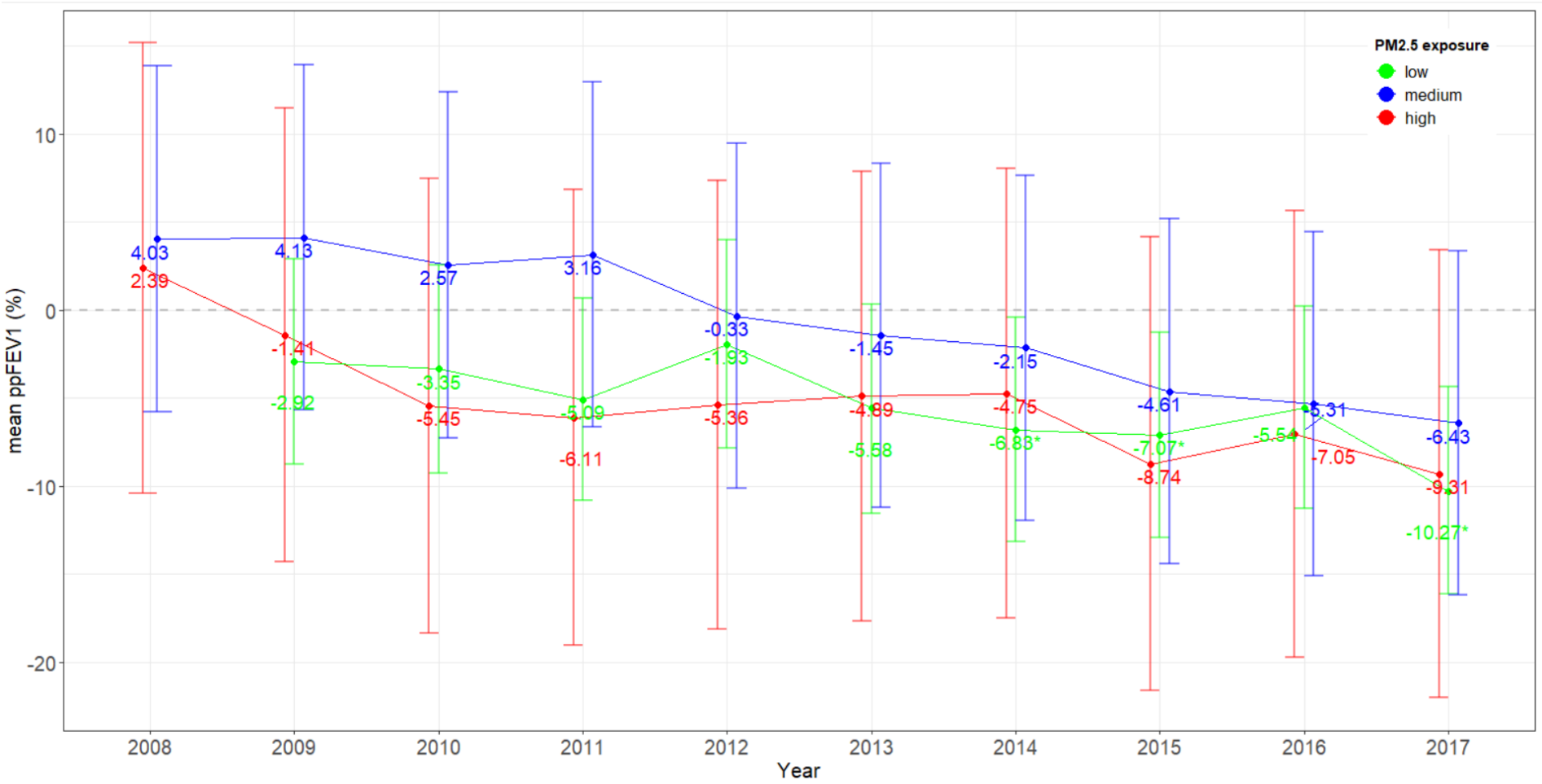


Figure 3.5: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF who did not change their residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

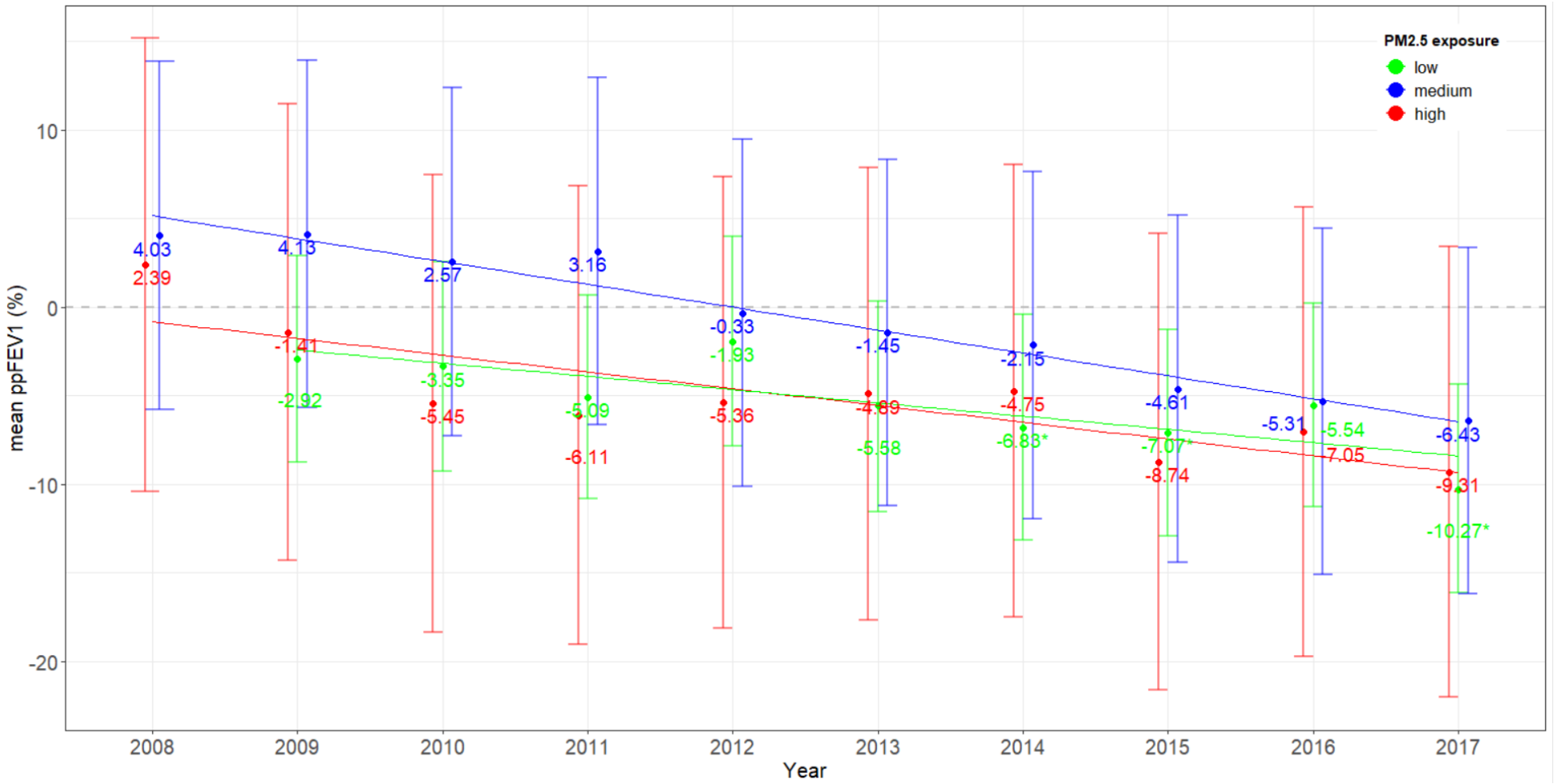


Figure 3.6: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of $PM_{2.5}$ among pwCF who did not change their residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

Table 3.9: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among children pwCF (6 – 15 years) with a residential postcode in the GLA between 2008 and 2017 (n= 122, ppFEV$_1$ measurements = 1,011)

		Crude			Adjusted[2]	
Characteristic	**Beta**	**95% CI[1]**	**p-value**	**Beta**	**95% CI[1]**	**p-value**
Interaction between year and PM$_{2.5}$ concentration (μg /m^3)						
2008 * low exposure	Ref	—		Ref	—	
2008 * middle exposure	1.39	-9.52, 12.3	0.802	2.05	-9.28, 13.4	0.722
2008 * high exposure	-10.4	-23.8, 3.01	0.128	-7.81	-21.2, 5.59	0.252
2009 * low exposure	-5.88	-13.8, 2.06	0.146	-7.98	-16.8, 0.86	0.077
2009 * middle exposure	-0.62	-11.5, 10.3	0.911	-0.13	-11.4, 11.2	0.981
2009 * high exposure	-6.32	-19.4, 6.80	0.344	-3.66	-16.7, 9.40	0.581
2010 * low exposure	-6.66	-14.4, 1.11	0.093	-7.67	-16.1, 0.76	0.074
2010 * middle exposure	-1.43	-12.3, 9.44	0.796	0.66	-10.7, 12.0	0.909
2010 * high exposure	-8.81	-21.9, 4.31	0.187	-6.43	-19.4, 6.57	0.331
2011 * low exposure	-8.81	-16.1, -1.53	**0.018**	-11.4	-19.3, -3.58	**0.004**
2011 * middle exposure	-2.01	-12.8, 8.83	0.716	-0.99	-12.3, 10.3	0.862
2011 * high exposure	-6.20	-19.2, 6.84	0.350	-3.11	-16.3, 10.1	0.642
2012 * low exposure	-9.74	-17.1, -2.36	**0.010**	-11.1	-19.1, -3.06	**0.007**
2012 * middle exposure	-4.03	-14.9, 6.80	0.464	-2.96	-14.2, 8.29	0.604
2012 * high exposure	-11.5	-24.6, 1.50	0.082	-9.28	-22.3, 3.76	0.162
2013 * low exposure	-8.97	-16.2, -1.70	**0.016**	-11.3	-19.2, -3.33	**0.005**
2013 * middle exposure	-5.32	-16.2, 5.51	0.334	-4.44	-15.7, 6.81	0.438
2013 * high exposure	-14.1	-27.2, -1.08	**0.034**	-11.5	-24.5, 1.54	0.084
2014 * low exposure	-11.0	-18.4, -3.60	**0.004**	-12.3	-20.6, -4.00	**0.004**
2014 * middle exposure	-4.25	-15.1, 6.57	0.440	-3.57	-14.8, 7.70	0.533
2014 * high exposure	-12.6	-25.6, 0.44	0.058	-10.5	-23.5, 2.51	0.113
2015 * low exposure	-12.4	-19.7, -5.17	**<0.001**	-12.8	-20.9, -4.83	**0.002**
2015 * middle exposure	-6.39	-17.2, 4.46	0.247	-5.27	-16.5, 6.01	0.358
2015 * high exposure	-16.3	-29.4, -3.27	**0.015**	-13.9	-26.8, -0.89	**0.036**
2016 * low exposure	-10.2	-17.5, -2.83	**0.007**	-11.8	-19.7, -3.94	**0.003**
2016 * middle exposure	-6.78	-17.6, 4.06	0.219	-5.73	-17.0, 5.53	0.317
2016 * high exposure	-15.0	-28.1, -1.99	**0.024**	-11.5	-24.5, 1.51	0.083
2017 * low exposure	-14.8	-22.1, -7.52	**<0.001**	-16.5	-25.0, -8.02	**<0.001**
2017 * middle exposure	-10.5	-21.4, 0.299	0.057	-8.99	-20.2, 2.27	0.117
2017 * high exposure	-18.6	-31.6, -5.64	**0.005**	-16.0	-28.9, -3.00	**0.016**

ppFEV$_1$ = percent predicted with FEV$_1$,

Low exposure of PM$_{2.5}$= pwCF who were exposed to either Q1 or Q2 of PM$_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

High exposure of PM$_{2.5}$= pwCF who were exposed to either Q4 or Q5 of PM$_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

Medium exposure of PM$_{2.5}$= all others who were not classified in either low or high exposure categories

[1]CI = Confidence Interval

[2]Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

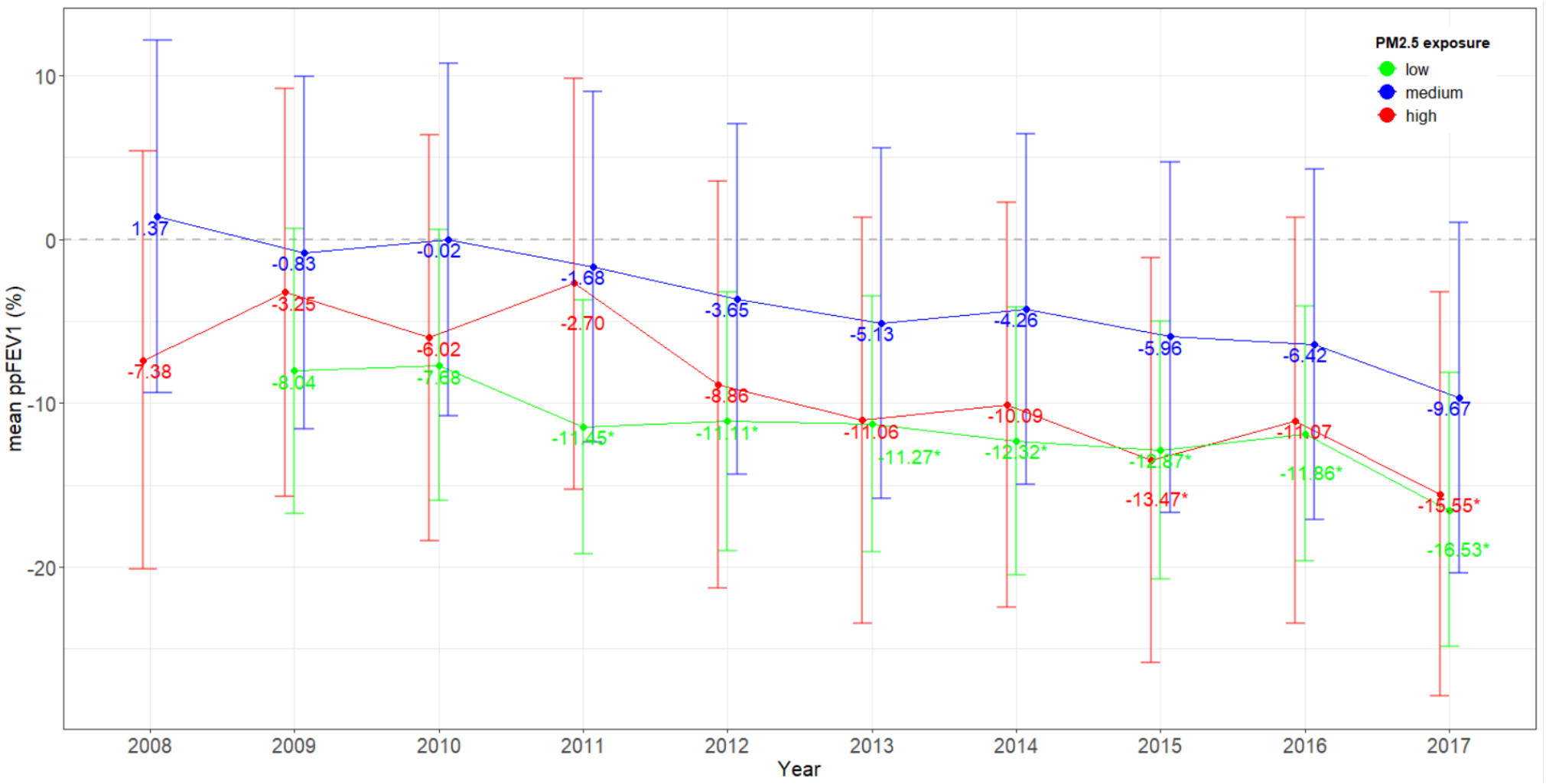


Figure 3.7: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among children pwCF (6 – 15 years) with a residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

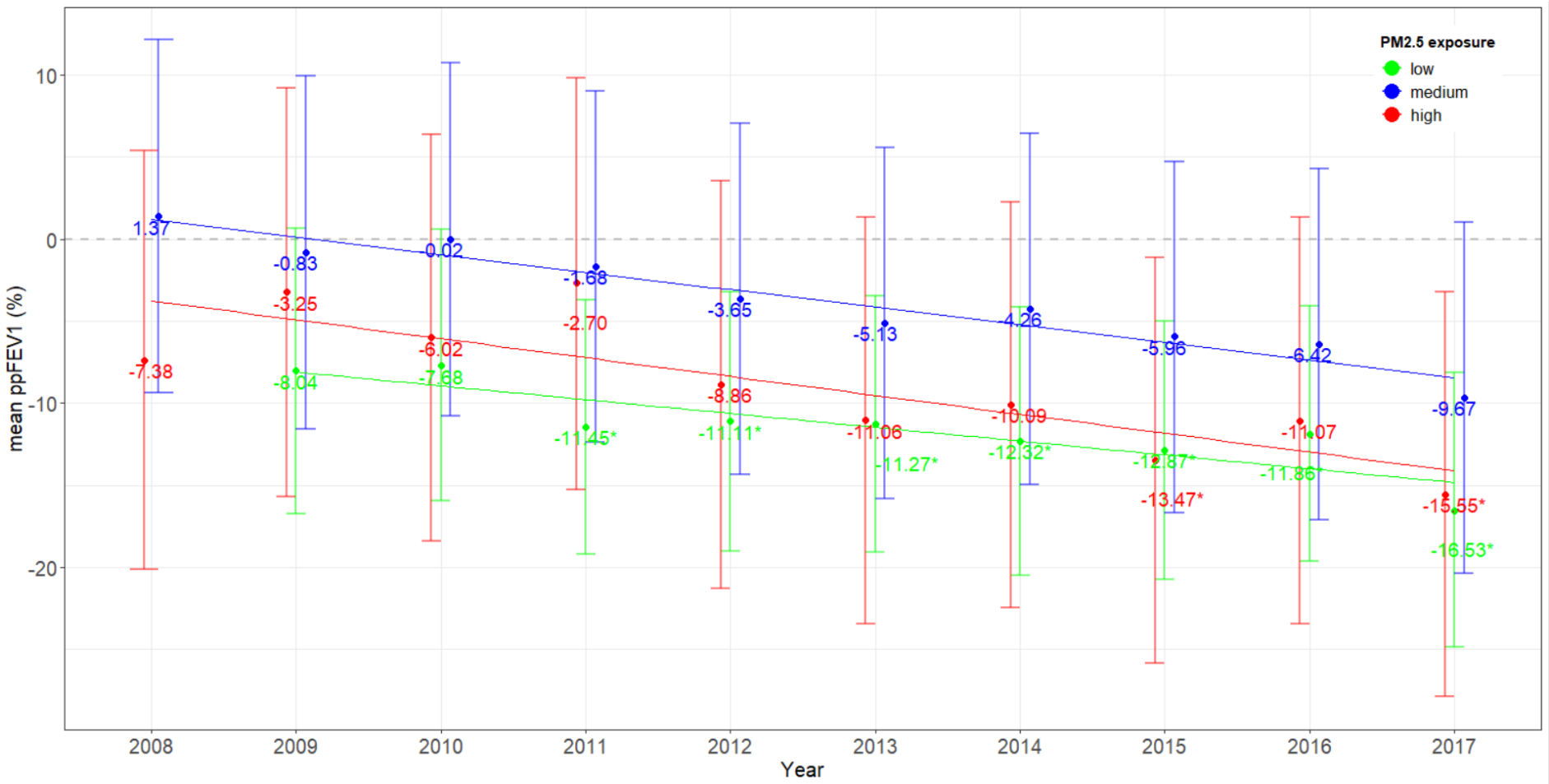


Figure 3.8: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of $PM_{2.5}$ among children pwCF (6 – 15 years) with a residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

Table 3.10: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among adults pwCF (≥16 years) with a residential postcode in the GLA between 2008 and 2017 (n= 271, ppFEV$_1$ measurements = 2,322)

		Crude			Adjusted[2]	
Characteristic	**Beta**	**95% CI[1]**	**p-value**	**Beta**	**95% CI[1]**	**p-value**
Interaction between year and PM$_{2.5}$ concentration (µg /m^3)						
2008 * low exposure	Ref	—		Ref	—	
2008 * middle exposure	-1.46	-9.02, 6.10	0.705	0.894	-6.23, 8.02	0.805
2008 * high exposure	-1.61	-11.7, 8.43	0.753	2.43	-7.03, 11.9	0.613
2009 * low exposure	-1.75	-5.21, 1.70	0.320	-1.07	-4.69, 2.55	0.562
2009 * middle exposure	-2.19	-9.74, 5.35	0.568	0.300	-6.81, 7.41	0.934
2009 * high exposure	-6.63	-16.6, 3.38	0.194	-2.41	-11.8, 7.02	0.615
2010 * low exposure	-1.53	-5.02, 1.96	0.389	-1.20	-5.04, 2.64	0.540
2010 * middle exposure	-3.59	-11.1, 3.94	0.349	-1.07	-8.19, 6.05	0.768
2010 * high exposure	-8.54	-18.6, 1.49	0.095	-4.36	-13.8, 5.13	0.367
2011 * low exposure	-2.67	-6.12, 0.778	0.129	-3.14	-6.82, 0.535	0.094
2011 * middle exposure	-3.65	-11.2, 3.88	0.341	-1.04	-8.15, 6.07	0.773
2011 * high exposure	-10.5	-20.5, -0.469	**0.040**	-6.19	-15.7, 3.31	0.201
2012 * low exposure	-0.493	-3.96, 2.97	0.780	-0.033	-3.73, 3.66	0.986
2012 * middle exposure	-6.43	-14.0, 1.10	0.094	-4.30	-11.4, 2.81	0.236
2012 * high exposure	-9.24	-19.2, 0.725	0.069	-5.15	-14.6, 4.27	0.283
2013 * low exposure	-2.79	-6.21, 0.633	0.110	-3.11	-6.78, 0.545	0.095
2013 * middle exposure	-7.30	-14.8, 0.231	0.057	-4.93	-12.0, 2.17	0.173
2013 * high exposure	-11.3	-21.3, -1.32	**0.027**	-7.74	-17.2, 1.72	0.109
2014 * low exposure	-2.77	-6.19, 0.651	0.112	-3.53	-7.23, 0.175	0.062
2014 * middle exposure	-7.65	-15.2, -0.120	**0.046**	-5.50	-12.6, 1.60	0.128
2014 * high exposure	-12.1	-22.1, -2.12	**0.018**	-8.17	-17.6, 1.25	0.089
2015 * low exposure	-6.58	-10.0, -3.12	**<0.001**	-6.11	-9.74, -2.48	**<0.001**
2015 * middle exposure	-9.85	-17.4, -2.32	**0.011**	-7.49	-14.6, -0.400	**0.038**
2015 * high exposure	-12.0	-22.0, -1.97	**0.019**	-8.33	-17.8, 1.10	0.083
2016 * low exposure	-5.07	-8.54, -1.60	**0.004**	-4.78	-8.40, -1.17	**0.010**
2016 * middle exposure	-10.9	-18.5, -3.41	**0.005**	-8.51	-15.6, -1.41	**0.019**
2016 * high exposure	-14.9	-24.9, -4.90	**0.004**	-10.9	-20.3, -1.44	**0.024**
2017 * low exposure	-5.37	-8.81, -1.94	**0.002**	-5.46	-9.15, -1.76	**0.004**
2017 * middle exposure	-12.1	-19.6, -4.54	**0.002**	-9.59	-16.7, -2.48	**0.008**
2017 * high exposure	-17.0	-27.0, -7.03	**<0.001**	-13.0	-22.4, -3.65	**0.007**

ppFEV$_1$ = percent predicted with FEV$_1$,

Low exposure of PM$_{2.5}$= pwCF who were exposed to either Q1 or Q2 of PM$_{2.5}$ concentration (µg /m^3) at least 5 times between 2008 and 2017 in London

High exposure of PM$_{2.5}$= pwCF who were exposed to either Q4 or Q5 of PM$_{2.5}$ concentration (µg /m^3) at least 5 times between 2008 and 2017 in London

Medium exposure of PM$_{2.5}$= all others who were not classified in either low or high exposure categories

[1]CI = Confidence Interval

[2]Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

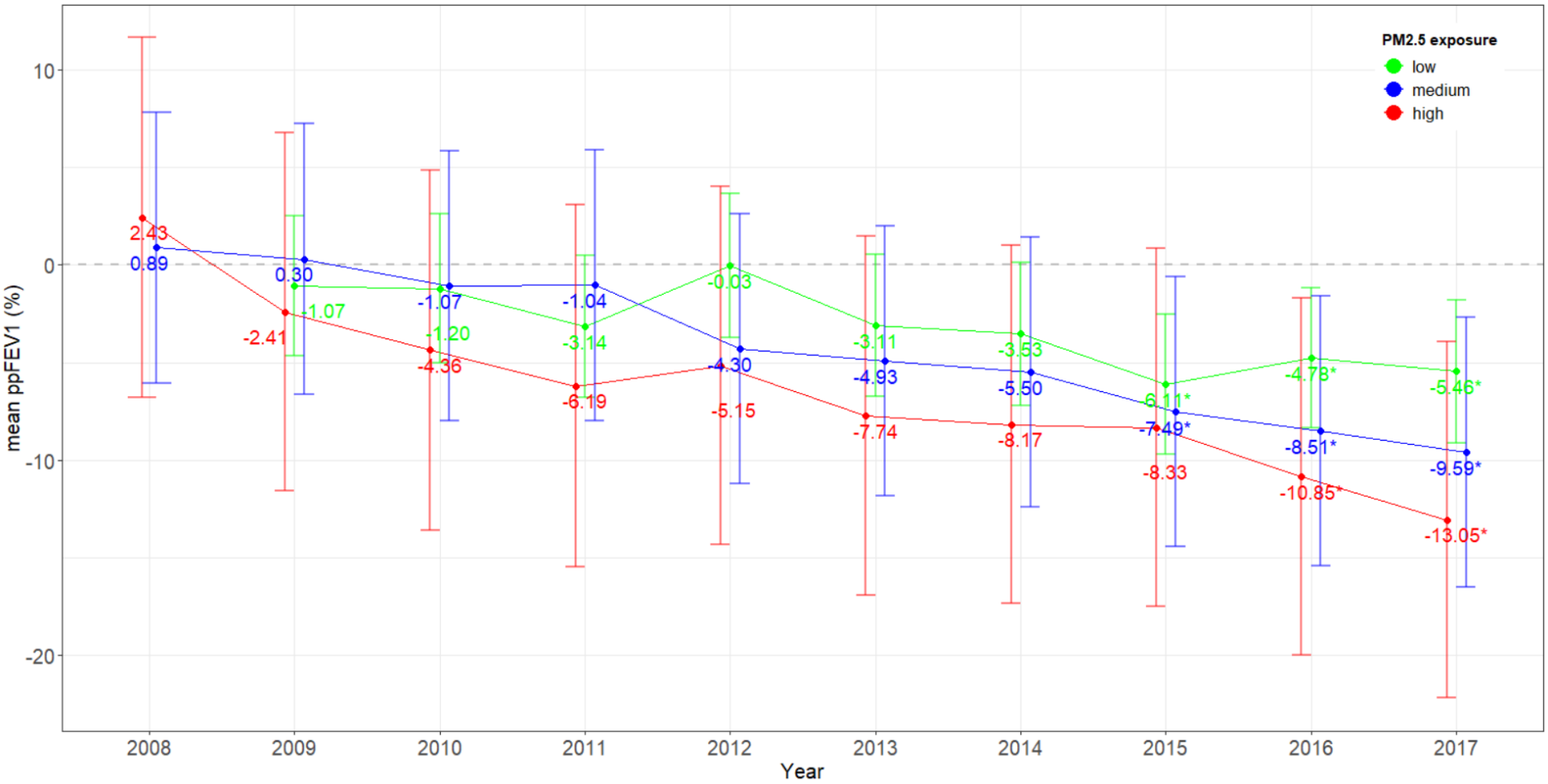


Figure 3.9: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among adults pwCF (≥16 years) with a residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

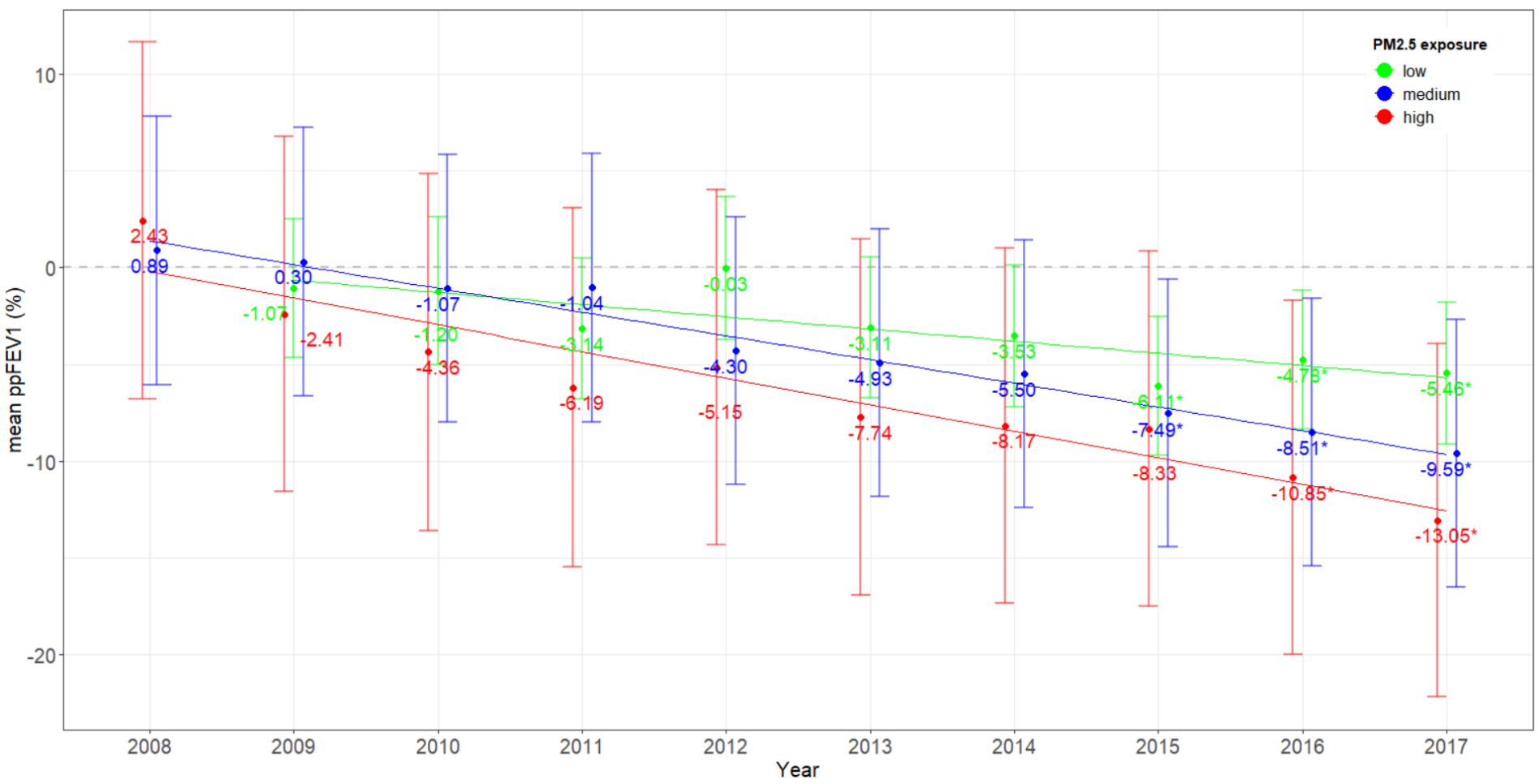


Figure 3.10: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among adults pwCF (≥16 years) with a residential postcode in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight, genotype *P. aeruginosa* and deprivation (Townsend index)

Table 3.11: Estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF with severe genotype (n= 302, ppFEV$_1$ measurements = 2,587)

		Crude			Adjusted[2]	
Characteristic	**Beta**	**95% CI[1]**	**p-value**	**Beta**	**95% CI[1]**	**p-value**
Interaction between year and PM$_{2.5}$ concentration (μg /m^3)						
2008 * low exposure	Ref	—		Ref	—	
2008 * middle exposure	3.56	-3.49, 10.6	0.321	2.21	-4.15, 8.56	0.496
2008 * high exposure	0.33	-8.60, 9.25	0.943	1.70	-6.38, 9.78	0.680
2009 * low exposure	-2.55	-6.33, 1.24	0.187	-1.78	-5.70, 2.13	0.372
2009 * middle exposure	2.55	-4.48, 9.59	0.476	1.28	-5.07, 7.62	0.693
2009 * high exposure	-1.79	-10.6, 7.04	0.690	-0.64	-8.62, 7.34	0.875
2010 * low exposure	-2.53	-6.30, 1.25	0.190	-1.86	-5.94, 2.22	0.371
2010 * middle exposure	0.67	-6.35, 7.69	0.851	-0.19	-6.54, 6.15	0.952
2010 * high exposure	-4.45	-13.3, 4.40	0.324	-3.54	-11.5, 4.45	0.385
2011 * low exposure	-5.80	-9.51, -2.08	**0.002**	-5.67	-9.52, -1.81	**0.004**
2011 * middle exposure	0.75	-6.26, 7.77	0.833	-0.25	-6.58, 6.08	0.939
2011 * high exposure	-4.32	-13.1, 4.49	0.336	-3.16	-11.2, 4.87	0.440
2012 * low exposure	-3.61	-7.32, 0.11	0.057	-2.95	-6.81, 0.92	0.135
2012 * middle exposure	-1.96	-8.97, 5.05	0.583	-3.37	-9.70, 2.97	0.297
2012 * high exposure	-5.53	-14.3, 3.27	0.217	-4.56	-12.5, 3.40	0.261
2013 * low exposure	-3.83	-7.52, -0.15	**0.041**	-3.98	-7.85, -0.11	**0.044**
2013 * middle exposure	-3.02	-10.0, 3.99	0.398	-4.29	-10.6, 2.03	0.183
2013 * high exposure	-7.46	-16.3, 1.34	0.097	-6.26	-14.3, 1.72	0.124
2014 * low exposure	-4.51	-8.19, -0.82	**0.017**	-4.99	-8.92, -1.07	**0.013**
2014 * middle exposure	-2.71	-9.72, 4.30	0.448	-4.29	-10.6, 2.03	0.183
2014 * high exposure	-7.32	-16.1, 1.48	0.103	-6.62	-14.6, 1.33	0.102
2015 * low exposure	-7.06	-10.8, -3.34	**<0.001**	-6.08	-9.94, -2.22	**0.002**
2015 * middle exposure	-5.78	-12.8, 1.24	0.107	-6.96	-13.3, -0.64	**0.031**
2015 * high exposure	-8.78	-17.6, 0.03	0.051	-8.09	-16.0, -0.15	**0.046**
2016 * low exposure	-5.29	-8.99, -1.59	**0.005**	-4.72	-8.51, -0.942	**0.014**
2016 * middle exposure	-5.92	-12.9, 1.10	0.098	-7.00	-13.3, -0.68	**0.030**
2016 * high exposure	-9.99	-18.8, -1.20	**0.026**	-8.55	-16.5, -0.63	**0.035**
2017 * low exposure	-7.32	-11.0, -3.63	**<0.001**	-6.61	-10.5, -2.67	**0.001**
2017 * middle exposure	-8.31	-15.3, -1.29	**0.020**	-9.28	-15.6, -2.96	**0.004**
2017 * high exposure	-12.8	-21.6, -3.99	**0.004**	-11.8	-19.7, -3.86	**0.004**

ppFEV$_1$ = percent predicted with FEV$_1$,

Low exposure of PM$_{2.5}$= pwCF who were exposed to either Q1 or Q2 of PM$_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

High exposure of PM$_{2.5}$= pwCF who were exposed to either Q4 or Q5 of PM$_{2.5}$ concentration (μg /m^3) at least 5 times between 2008 and 2017 in London

Medium exposure of PM$_{2.5}$= all others who were not classified in either low or high exposure categories

Severe genotype = homozygous *F508del*, *F508del*/Minimal function (*MF*), *Gating* and *MF/MF*

[1]CI = Confidence Interval

[2]Adjusted for age at baseline, sex, height weight, *P. aeruginosa,* and deprivation (Townsend index)

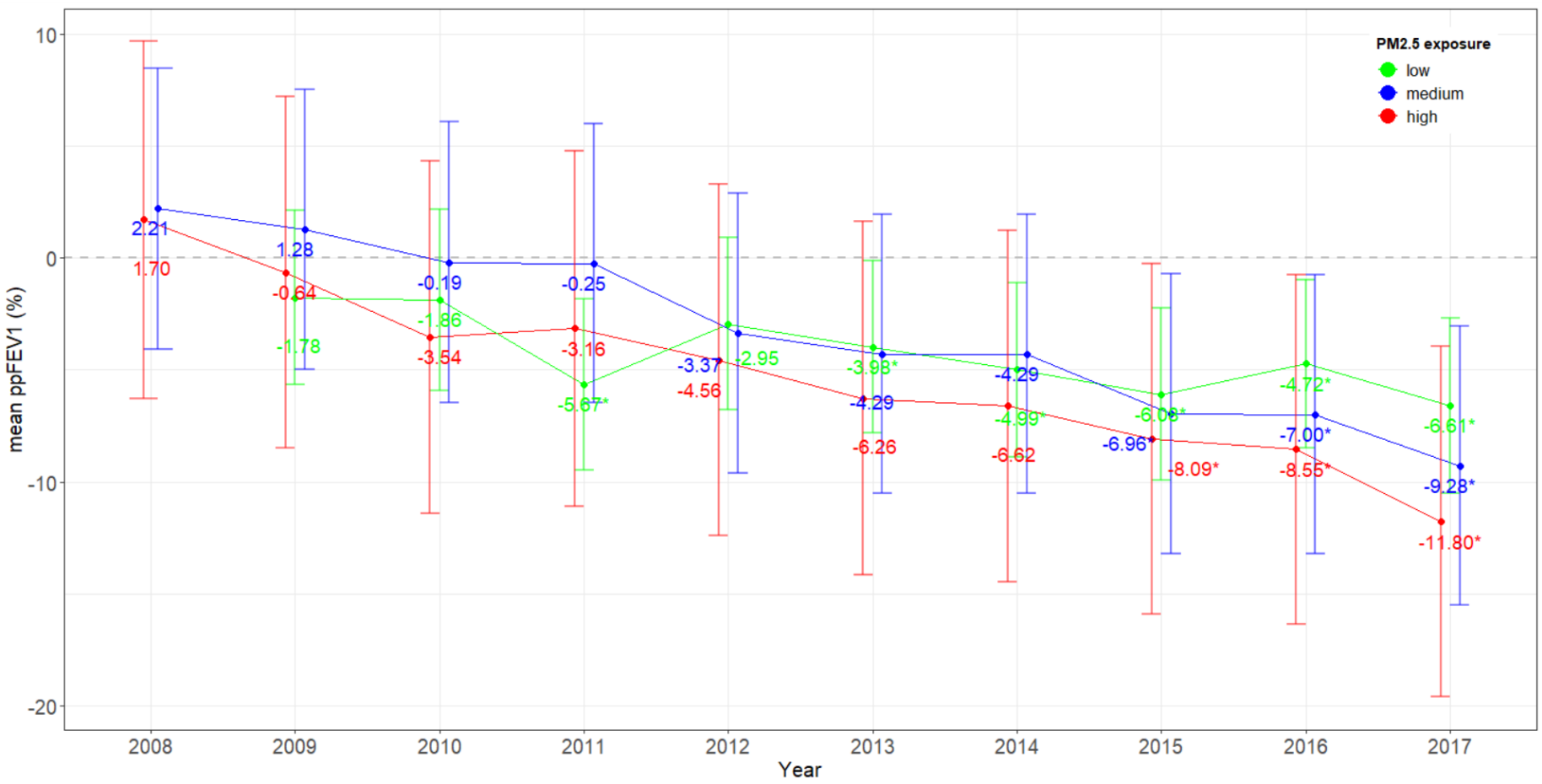


Figure 3.11: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of $PM_{2.5}$ among pwCF with a severe genotype in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight *P. aeruginosa,* and deprivation (Townsend index)

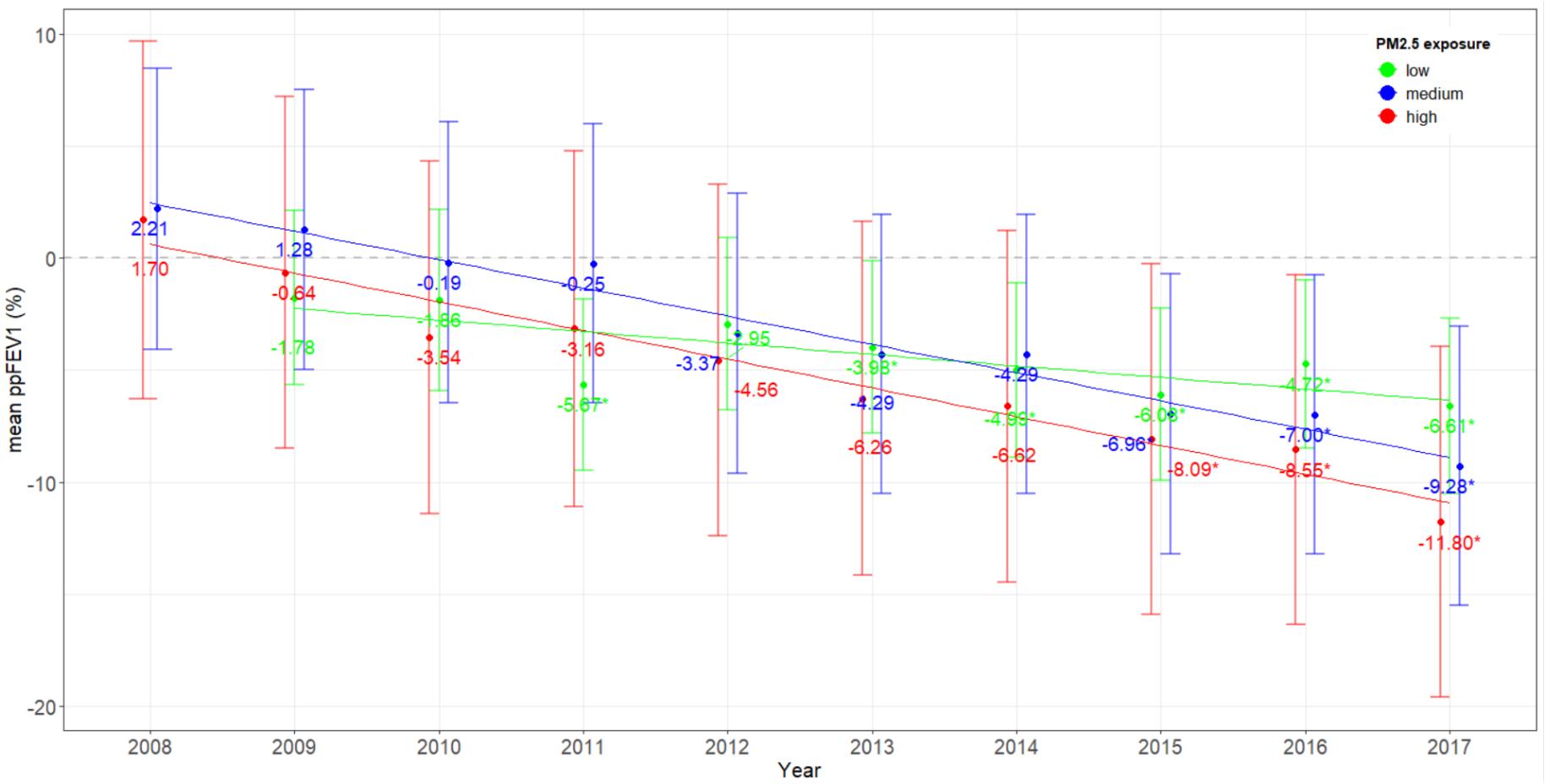


Figure 3.12: Adjusted* estimated change in mean ppFEV$_1$ associated with annual exposure of PM$_{2.5}$ among pwCF with a severe genotype in the GLA between 2008 and 2017

*Adjusted for age at baseline, sex, height, weight *P. aeruginosa,* and deprivation (Townsend index)

3.4 Discussion

This study aimed to explore the effects of ambient $PM_{2.5}$ exposure on $ppFEV_1$ decline in pwCF across the GLA over a ten-year period, from 2008 to 2017. This study used high-quality data from the UK CF Registry of all pwCF with a residential postcode within the GLA and used a high-resolution modelled surface of $PM_{2.5}$ using the "London Air Pollution Toolkit". In my analysis, I explored three potential timings for $PM_{2.5}$ exposure: annual average, weekly average and one-day average before the annual review date. Given my hypothesis focusing on long-term cumulative exposure and aligned with the yearly frequency of outcome review in the Registry. Additionally, as demonstrated in ***Figure 3.3***, the annual average exposure exhibited less variation compared to the other two timeframes. Therefore, I chose to present results using the annual average exposure. This approach ensures consistency with my research question and maximizes the utilization of available data. While the other timing options offer valuable insights for potential future investigations, the annual average best addresses the current research objective. In addition, the linear mixed models' analysis took into consideration the age, sex, height, weight, genotype of *P. aeruginosa* sputum culture positive and deprivation (Townsend index). The results of this study highlight the significant relationship between higher $PM_{2.5}$ exposure and the decline in $ppFEV_1$ trajectories among pwCF with a Greater London residential postcode. These results were also consistent in the stratified subgroup analysis of CF adults. The findings of this study are novel as there have been no prior longitudinal studies with a repeated measure to demonstrate the effect of $PM_{2.5}$ exposure on changes in $ppFEV_1$ trajectories over time in pwCF. This work was inspired by a conversation with environmental research specialists Ian Mudway and Sean Beavers, and the fact that changes in $ppFEV_1$ are a better marker of disease progression than individual measurements. Therefore, the focus of this analysis was on the change in $ppFEV_1$ over time. The analysis's main emphasis was how $ppFEV_1$ changed over time. To understand more about how $PM_{2.5}$ affects lung function trajectories over time in pwCF, I have built on my prior study (key chapter I of this thesis), which revealed a link between $PM_{2.5}$ and $ppFEV_1$ in a semi-individual cross-sectional research design in pwCF in the UK CF Registry.

3.4.1 Lung function ($ppFEV_1$)

$ppFEV_1$ is a key prognostic marker in pwCF (Earnest *et al.*, 2020). The decline in $ppFEV_1$ over time plays an important role in monitoring CF disease progression, guiding treatment based on pwCF responses to therapy, and improving overall clinical outcomes in pwCF. In the current analysis, overall, there was a 1.08% (95%CI: -1.21; -0.95) annual average decline in $ppFEV_1$ among pwCF with a residential postcode in London. This implies that pwCF in Greater London experiences a gradual decline in $ppFEV_1$ over time.

Comparing my finding with another study which used the same UK CF Registry data from 2015 to 2017 found that CF adults (n = 5,554) had a slightly higher decline in $ppFEV_1$ (-1.52% (95% CI: -1.66; -1.38) (Caley *et al.*, 2021)). They had a significantly larger sample size, but they only used repeated $ppFEV_1$ measurements taken over two years, whereas I used repeated $ppFEV_1$ measurements taken over ten years for my study. By investigating across a longer time frame, it may be possible to get a precise estimate of the average yearly reduction in $ppFEV_1$ while simultaneously taking into consideration any potential variations and fluctuations in lung function. Furthermore, another study with a larger sample size of 32,843 pwCF found that the average annual decline in $ppFEV_1$ was 1.98% (95% CI: 1.86; 2.10), which is almost twice as much as what I found (Szczesniak *et al.*, 2013). This substantial difference in the decline rate at the population level highlights the potential variability in the $ppFEV_1$ decline.

3.4.2 Lung function ($ppFEV_1$) and $PM_{2.5}$

$PM_{2.5}$ is one of the main air pollutants with a substantial effect on respiratory health. There has been substantial research on how $PM_{2.5}$ affects lung function in the general population and people with non-CF respiratory disease, which supports the association between high concentrations of $PM_{2.5}$ levels and a higher risk of respiratory symptoms, exacerbation of respiratory conditions, and the development of respiratory diseases (Bălă *et al.*, 2021; Hoffmann *et al.*, 2022; Kyung & Jeong, 2020). Therefore, understanding the effects of $PM_{2.5}$ on lung function decline in vulnerable populations, such as those with CF, is essential for implementing effective strategies to improve the overall health of pwCF. In line with my hypothesis, this longitudinal analysis revealed a significant association between high $PM_{2.5}$ levels and a decline in $ppFEV_1$ trajectories in pwCF. This association between high $PM_{2.5}$

exposure and a decline in ppFEV$_1$ over time remained significant even after adjusting for various potential confounders, including age, sex, genotype, height, weight, *P. aeruginosa* and deprivation, suggesting that exposure to high levels of $PM_{2.5}$ concentration has an independent effect on the ppFEV$_1$ decline in pwCF.

To the best of my knowledge, the association between $PM_{2.5}$ and a decline in ppFEV$_1$ in the CF population has not been investigated before in a longitudinal study. However, there is one study with a large sample size (n = 11,484) that examined the relationship between $PM_{2.5}$ and lung function in pwCF in a cohort study (Goss *et al.*, 2004). I used quintiles to assess the $PM_{2.5}$ exposure in this analysis, whereas Goss *et al.* employed a continuous approach. Despite these methodological differences in the $PM_{2.5}$ exposure measurement, both investigations found that $PM_{2.5}$ had a negative effect on pwCFs' lung function. The findings from Goss *et al.* showed a 0.5% (95% CI: 0.3; 0.9%) reduction in FEV$_1$ for every 10 µg/m3 increase in $PM_{2.5}$. Comparing the findings of the two studies, it is notable that the effect size reported by Goss *et al.* is slightly higher than the effect observed in my study. However, it is important to consider the differences in the approach to analysing $PM_{2.5}$ exposure in both studies. The use of quintiles in my study allows for the exploration of the impact of $PM_{2.5}$ across different exposure levels, providing a broader perspective on the relationship between $PM_{2.5}$ exposure and lung function in pwCF. The findings of both investigations demonstrate the detrimental effect that $PM_{2.5}$ has on CF patients' lung function. These findings highlight how important it is to make efforts to reduce $PM_{2.5}$ exposure and improve air quality, particularly for pwCF who are already at risk for respiratory complications.

Using a longitudinal study design with repeated ppFEV$_1$ measurements over time could help us better understand the prognosis of CF disease, allow us to explore temporal relationships between trajectories of $PM_{2.5}$ and ppFEV$_1$ and enable us to identify causal relationships that can be used to develop targeted treatments and preventive measures. The results of this study contribute to the growing body of evidence of an association between $PM_{2.5}$ and ppFEV$_1$. This study also highlights the importance of identifying and addressing pwCF environment-specific vulnerabilities. The implications of the results are important from both clinical and public health perspectives. The findings of this study stress the importance of lowering exposure to high levels of $PM_{2.5}$ concentration to preserve and improve lung function in pwCF This may be accomplished by raising knowledge of different guidelines,

such as the air quality index in London and the WHO legal limit of $PM_{2.5}$, which was recently revised to 5 µg/m^3, and avoiding densely populated places during periods of high pollution by following the air quality index in London.

I will compare the findings of this longitudinal analysis with research on other respiratory disorders and the general population to understand the broader impact of $PM_{2.5}$ on ppFEV_1 trajectories. A prospective large sample study that included 1,759 schoolchildren (mean age 10 years) to examine the effect of air pollution on lung function reported that children exposed to high levels of $PM_{2.5}$ had significantly lower FEV_1 (7.9% vs. 1.6% p = 0.002) than children exposed to low levels of $PM_{2.5}$ (Gauderman *et al.*, 2004). Another birth cohort study that included 915 children from the GINIplus and LISA birth cohorts in Germany to examine the relationship between air pollution exposure during the first year of life and lung function up to the age of 15 years reported an annual decline in FEV_1 z-scores of -0.012 (95% CI: -0.014; -0.009) for every interquartile range (1.1 µg/m^3) increase in $PM_{2.5}$ levels (Zhao *et al.*, 2021). Further, a longitudinal analysis of 915 children from the Dutch PIAMA birth cohort research found that FEV_1 decreased by 0.26% (95%CI: - 0.49; - 0.03%) per year for every 1.2 µg/m^3 increase in $PM_{2.5}$ (Milanzi *et al.*, 2018). Zhang *et al.* recently published a meta-analysis to investigate the adverse effects of acute $PM_{2.5}$ exposure on lung function in healthy and asthmatic children. The results showed that the adverse effects of $PM_{2.5}$ on lung function were more severe in children with asthma (3.11, 95% CI: 4.54, 1.67) than they were in healthy children (1.61, 95% CI: 2.34; 0.91) (Zhang *et al.*, 2023). Whereas in an adult longitudinal study, 133,119 individuals (18 years or above) were recruited between 2001 and 2014 in Taiwan to investigate the associations between long-term exposure to ambient $PM_{2.5}$, changes in lung function and the incidence of COPD. The difference in FEV_1 between individuals with high and low exposure to $PM_{2.5}$ was 7.66 mL (9.93, 5.41)/year (Bo *et al.*, 2021). This comparison with other studies showed a consistent pattern between pwCF and other respiratory diseases, indicating that $PM_{2.5}$ has a significant adverse effect on lung function decline across different populations. These findings also highlight the importance of understanding the broader implications of exposure to $PM_{2.5}$ on respiratory health beyond CF. It highlights the need for comprehensive air pollution control measures and public health interventions to reduce the impact of $PM_{2.5}$ on lung function trajectories.

These results underline the need for patient-specific care plans and emphasise the importance of continued research to better understand the factors that are significantly associated with the decline of FEV_1 trajectories in pwCF. Understanding the $ppFEV_1$ trajectories in pwCF is important, as early identification of poor $ppFEV_1$ trajectories will help to improve lung function and delay the course of the disease progression. These FEV_1 trajectories are influenced by a variety of factors, including age, gender, height, genotype type, co-morbidities, individual response to treatment, and environmental exposures. pwCF with a rapid decline in FEV_1 are associated with increased worsening of respiratory symptoms, a high healthcare burden due to increased use of the healthcare system, and poor survival rates. Several research studies have explored the factors associated with the decline of $ppFEV_1$ trajectories (Earnest *et al.*, 2020; Hatziagorou *et al.*, 2023; Taylor-Robinson *et al.*, 2012); however, there is limited data on the effects of $PM_{2.5}$ exposure on $ppFEV_1$ trajectories in longitudinal investigations in pwCF. A better understanding of the long-term effects of $PM_{2.5}$ on the decline in $ppFEV_1$ trajectories over time in pwCF is imperative for preventive and management strategies to reduce the burden of the impact of $PM_{2.5}$ on $ppFEV_1$ trajectories in the CF population.

I found a comparably low percentage of *P. aeruginosa* culture positives in this analysis cohort and across all pwCF in the UK CF registry in 2008 and 2009. This low prevalence could be due to sampling error or reporting error. In 2016 and 2017, across all pwCF in the UK CF registry cohort, we also found a substantial reduction in positive *P. aeruginosa* cultures, which was not seen in the GLA cohort. The reason for this is unclear as we know of no differences in clinic segregation, infection monitoring, and *P. aeruginosa* eradication therapy policies during this time Other observational studies also found a decreasing trend in the prevalence of *P. aeruginosa* infection in pwCF (Ramsay *et al.*, 2017; Salsgiver *et al.*, 2016). Further research is required to determine if this decline in culture-positive *P. aeruginosa* persists and how it correlates with lung function trajectories in pwCF.

In my analysis of the Greater London cohort, the most deprived quintile had the higher prevalence of pwCF patients (43.5%), while the least deprived quintile had the smallest number of patients (5.9%). One possible reason for this difference could be the small number of pwCF with residential postcodes in Greater London. There was an almost equal distribution of pwCF across all Townsend index quintiles in the UK CF registry cohort.

However, this distribution points to a possible relationship between socioeconomic disadvantage and CF prevalence, with a larger burden of the CF disease shown in regions with higher degrees of deprivation. A number of studies have consistently reported that people from the most deprived area experience worse health outcomes than those in the least deprived area (Taylor-Robinson *et al.*, 2015; Taylor-Robinson *et al.*, 2013). To understand why more pwCF patients may be found in the most deprived area in Greater London compared to the rest of the UK, more research focusing on CF prevalence in other areas would be warranted.

3.4.3 Sensitivity analyses

3.4.3.1 Among non-movers

Contrary to my expectations, there were no clear signals of associations between $PM_{2.5}$ exposure levels and $ppFEV_1$ trajectories in the subgroup of pwCF who did not change their residential postcode during the 2008 – 2017 study period. It is unclear why there is a weak association in this subgroup, but we discuss below possible explanations. First, this could be the small sample size, which would reduce statistical power and make it difficult to detect associations between $PM_{2.5}$ exposure levels and $ppFEV_1$ trajectories that may exist in this subset of pwCF (Serdar *et al.*, 2021). Another possible explanation is the influence of healthy study participants; pwCF who did not move may represent a population that is relatively healthier than pwCF who did as pwCF with a severe disease may move and relocate to have better access to care or move to an area where the $PM_{2.5}$ concentrations are low.

There were some variations in mean $ppFEV_1$ among pwCF with medium and high exposure to $PM_{2.5}$ between 2009 and 2012 ***(Figure 3.5)***, which reduced after that year. I am not sure of any probable causes, but further in-depth analysis will help identify plausible causes. This includes the frequency distribution of plots by $PM_{2.5}$ exposure levels and by year, as well as a subgroup analysis of these years. To completely comprehend the underlying mechanisms and factors that may impact the association between $PM_{2.5}$ exposure and $ppFEV_1$ trajectories in this subset of pwCF who did not change postcodes, more studies with a larger sample size are required.

3.4.3.2 Among children and adults

We found strong signals of a relationship between levels of $PM_{2.5}$ exposure and declines in $ppFEV_1$ trajectories in CF adults but not in CF children. Our cohort of children was also smaller than that of adults, which likely affected the power of this sub-group analysis to identify any significant association. Children's developing lungs have a strong ability for repair and regeneration as they grow, which may mask the impact of $PM_{2.5}$ on lung function in children. Furthermore, CF Children spend a lot of time indoors, where the level of $PM_{2.5}$ they are exposed to may depend on the quality of the indoor air. PEx, which are associated with rapid $ppFEV_1$ decline, tend to be more common in adults than in children. Overall, the absence of a clear association between $PM_{2.5}$ exposure level and $ppFEV_1$ in CF children highlights the need for more research studies to understand the mechanisms that underline and associate factors. My analysis may scale up to whole GB or level with relevant exposure data to better understand the complex relationship between $PM_{2.5}$ exposure and $ppFEV_1$ trajectories in CF children. These findings underscore the importance of considering individual variability and age-specific characteristics when exploring the impact of $PM_{2.5}$ exposure on pulmonary health in CF children. There has been some inconsistency in the findings of earlier investigations on the impact of $PM_{2.5}$ on lung function in children. There is evidence from several earlier studies on asthma (Qibin *et al.*, 2020; Sbihi *et al.*, 2017), COPD (Bo *et al.*, 2021; Li *et al.*, 2022) and others (Yang *et al.*, 2020; Zhang *et al.*, 2023; Zhao *et al.*, 2021) found a significant association between $PM_{2.5}$ exposure and children's FEV_1 decline trajectories however, other studies have shown no association between $PM_{2.5}$ exposure and FEV_1 decline in children (Delfino *et al.*, 2008; Takebayashi *et al.*, 2022).

The inconsistency in these studies may be attributed to differences in study design, sample size, study location, exposure assessment tools and methods, and the adjustment of potentially confounding covariates in the regression model. More research studies are necessary to understand the underlying mechanisms and factors contributing to the differences in these findings.

3.4.3.3 Among severe genotypes

There was a significant association between $PM_{2.5}$ exposure level and $ppFEV_1$ among the subset of CF with a severe genotype. These findings imply that the association that I found

between $PM_{2.5}$ exposure levels and declines in $ppFEV_1$ in pwCF with severe disease is reliable and robust. pwCF with severe genotype tend to be sicker and therefore are more likely to have a worse decline in ppFEV1 (Chen *et al.*, 2021; Mei Zahav *et al.*, 2023). This finding of a significant association in the subset group suggests that $PM_{2.5}$ exposure has a measurable effect on FEV_1 in pwCF with a severe genotype. This supports the growing body of evidence highlighting the adverse effects of $PM_{2.5}$ on lung function, especially in vulnerable populations like those with CF. The implications of these results highlight the importance of reducing $PM_{2.5}$ exposure levels in pwCF with a severe genotype. Additionally, these results highlight the potential need for targeted and more intensive interventions for pwCF who have a severe genotype and live in areas with high and medium exposure to $PM_{2.5}$. More research is recommended by genotype to validate this finding.

This severe genotype is based on the new classification that I developed in this project so there is no direct comparison of this finding with other studies. However, there are two other studies carried out in a subgroup of pwCF homozygous for the *F508* mutation. First, a study from the Canadian CF Registry by Leung *et al.* found a decline of 1.22% (95% CI: 1.33;1.13) in the subset of 1,789 homozygous *F508* pwCF between 2007 and 2016 (Leung *et al.*, 2020). The second study was conducted by Sawicki et al., Using data from 2006 to 2014 from the US CF Foundation Patient Registry, Sawicki *et al.* found a decline of 1.91% (95% CI: -2.01; -1.80) in $ppFEV_1$ per year in *F508del* homozygous cohorts of 11,916 pwCF with a mean (SD) age of 18.0 (9.6) years and a 0.70% (95% CI: -1.09, -0.30) decline in the *F508del* /RF cohort of 1,242 pwCF with a mean (SD) age of 23.0 (12.1) years (Sawicki *et al.*, 2022). The results of the two studies were based on homozygous *F508del* pwCF, but my research sample covered all pwCF. These results provide valuable insights into the heterogeneity of the disease and may call for more research in future studies by focusing on a genotype subgroup.

3.4.4 Strengths of the study

To the best of my knowledge, this is the first longitudinal research using repeated measurements to assess the relationship between long-term $PM_{2.5}$ exposure and the decline in $ppFEV_1$ over 10 years in pwCF. A longitudinal study with repeated measurements has the main benefit of establishing a temporal relationship between $PM_{2.5}$ exposure and FEV_1 decline in pwCF also changes in $ppFEV_1$ over time are a better measure of disease

progression than an individual measurement of $ppFEV_1$. This study's key strength was the availability of repeated measurements of $ppFEV_1$ and comprehensive individual postcode level $PM_{2.5}$ exposure data from a high-resolution map for the study period, which allowed me to examine the exposure's long-term effects on pwCF's $ppFEV_1$ decline. Additionally, by collecting data at multiple points in time, I was able to establish the temporal relationship between $PM_{2.5}$ exposure and FEV_1 decline. The longitudinal research design also allows for assessing the effect of $PM_{2.5}$ exposure on $ppFEV_1$ decline over time and provides cumulative effects between $PM_{2.5}$ exposure and $ppFEV_1$ decline in pwCF. Additionally, using a longitudinal research design improved statistical strength by reducing the likelihood of within pwCF's variation over time and improving the statistical power to detect the association if there is any. This study design also helps to control time-variant confounding variables like age, height, weight, and number of *P. aeruginosa* cultures, which helps to reduce the confounding bias, strengthen the study's internal validity, and increase the ability to attribute observed effects to $PM_{2.5}$ exposure. The use of standardised approaches and methods for $ppFEV_1$ values using the Global Lung Function Initiative (GLI) prediction equation was also one of the study's strengths. (Quanjer *et al.*, 2012).

3.4.5 Limitations of the study

My study is not without limitations. Outdoor $PM_{2.5}$ concentrations may not fully reflect pwCF exposure as pwCF may spend significant periods indoors. Therefore, the results should be interpreted in the context of the absence of this exposure that was not considered in this study, such as indoors, in a school, or at work, which may reduce the precision of the $PM_{2.5}$ exposure estimation. However, there are many previous studies which have used residential exposure to assess the health risk (Ashworth *et al.*, 2021; Hajna *et al.*, 2023; Ronaldson *et al.*, 2022). There is a growing focus on indoor air quality and exposure in different places, but indoor exposure tends to be associated with outdoor exposure. The exposure level of $PM_{2.5}$ was assigned to the pwCF residential postcode. Information on the patterns of daily activities of pwCF was not available. More advanced technologies are needed to assess the pwCF individual exposure more accurately in future studies. Nevertheless, the current study adds a novel perspective to the literature on the long-term effects of $PM_{2.5}$ on $ppFEV_1$ decline in pwCF in Greater London. A small sample size in the subgroup of analysis might have reduced the power of the study in stratified analysis. Smoking, pancreatic insufficiency,

dietary condition, and chronic pseudomonas infection are other potential contributing factors that were not available for this project. These factors are discussed in more detail in ***Section 2.4.3*** above. Longitudinal studies may provide valuable information about the association between exposure to $PM_{2.5}$ and a decline in $ppFEV_1$. There may still be residual confounding despite efforts to adjust for all potential confounders, unmeasured and unaccounted variables may influence the observed associations between $PM_{2.5}$ exposure and $ppFEV_1$ trajectories.

3.4.6 Conclusions

In conclusion, this study found that higher $PM_{2.5}$ exposure levels were adversely associated with a greater decline in $ppFEV_1$ in pwCF in GLA from 2008 to 2017. These findings contribute to the limited knowledge of the impact of exposure to $PM_{2.5}$ on $ppFEV_1$ decline in pwCF in London. Additional research with a larger sample size and a focus on indoor $PM_{2.5}$ exposure is needed to support the findings of this study and comprehend the effects of extensive exposure to $PM_{2.5}$ in pwCF.

4 Chapter IV: Bioaerosol exposure and disease severity in the UK

I published this work in the *Environmental Health* journal during the 3rd year of my PhD. The full citation of the full paper is doi: 10.1186/s12940-022-00932-1 Assessing the health risk of living near composting facilities on lung health, fungal and bacterial disease in cystic fibrosis: a UK CF Registry study | Environmental Health | Full Text (biomedcentral.com).

Khan, M.S., Douglas, P., Hansell, A.L. *et al.* Assessing the health risk of living near composting facilities on lung health, fungal and bacterial disease in cystic fibrosis: a UK CF Registry study. *Environ Health* **21**, 130 (2022). https://doi.org/10.1186/s12940-022-00932-1



4.1 Background

Waste disposal is one of the growing environmental health issues and a global challenge. It is estimated that every year over two billion metric tonnes of waste are produced globally, and one-third of this is not disposed of in an environmentally safe way (The World Bank., 2018). It is predicted that the amount of solid waste production will increase by threefold by 2100 worldwide, and the estimated cost to manage the waste will be over $350 billion per year by 2025 globally (Population matters., 2018). Every year, almost 330 million metric tonnes of waste are generated in the UK. Of the 221 million metric tonnes of waste produced in 2016 in the UK, 85% came from England (DEFRA., 2022).

Rapid growth in population and urbanisation are attributed to rising waste production; hence, the number of landfill sites in European countries is decreasing, therefore, it is important to manage the waste appropriately. The rise in waste resulted in the introduction of the "Union landfill directive (1999/31/EC)", which helped in the opening of new composting sites across European countries (European Commission., 2012). An operating permit is required in England for composting facilities that handle more than 60 metric tonnes of compost at once (DEFRA., 2014). In the UK, there are also more composting facilities than ever before (Robertson *et al.*, 2019). In England, there were 140 large-scale composting sites in 2010; by 2017, there were more than 300 large-scale composting plants (Robertson *et al.*, 2019).

Composting is one of the most effective ways of managing biodegradable waste, as it reduces the waste going to landfills and produces valuable fertility soil. The composting process could be done in open-air windrows or in-vessel systems (enclosed). Over two-thirds (~70%) of composting facilities are open windrows in the UK (Environment Agency., 2018). The three stages of composting are sanitisation, stabilisation and maturation. The agitation stage of composting, which includes turning, shredding, and screening, produces larger concentrations of bioaerosols (Taha *et al.*, 2006). The whole process of composting should be less hazardous to human health. Composting, however, poses serious concerns for public health because of the airborne bioaerosols that can be produced and emitted during this process (Douwes *et al.*, 2003). The emissions of bioaerosol from open composting facilities

are more hazardous than those from closed composting facilities, as their bioaerosol is released directly into the surrounding environment.

The bioaerosol may travel over long distances within the vicinity of these composting sites based on particle size, humidity level, wind speed and wind direction. Large particles over 30 micrometres in diameter can travel up to 100 metres from the source, medium-size particles between 10 and 30 micrometres in diameter can travel up to 500 meters, and smaller particles less than 10 micrometres may travel up to 1km (Clauss, 2015). Evidence suggests that some bioaerosols may not ground across lengths of a few kilometres from a PCS, although prior studies have demonstrated that bioaerosol concentrations falling within 250 metres are acceptable by the risk assessment of the UK Environment Agency (Environment Agency., 2016; Williams *et al.*, 2019).

The particles in these bioaerosols range in diameter from 0.2 to 100 µm and include fungi and fungal spores, both gram-positive and gram-negative bacteria that generate spores (such as actinomycetes), endotoxins, and other substances (Douwes *et al.*, 2003). Exposure to bioaerosols, including fungi, bacteria, and particulars, especially those smaller than 10 microns, has a harmful impact on human health, especially the respiratory system since these bioaerosols enter the alveolar sac of the lungs and trigger their function, which results in a negative health effect. Those who live close to major composting sites may be at greater risk for these health issues. There are various health effects related to these bioaerosols reported by Lacey *et al.* (Lacey & Dutkiewicz, 1994). These health effects include allergy bronchopulmonary aspergillosis (ABPA), rhinitis, bronchitis, allergic bronchitis, hypersensitivity pneumonitis, toxic pneumonitis, allergic asthma, chronic obstructive pulmonary disease (COPD), organic dust toxic syndrome, eye and skin irritations.

Although there is limited evidence of the association between bioaerosol emissions from composting sites and health effects in the neighbourhood, the potential risk cannot be completely disregarded (Douglas *et al.*, 2016a; Robertson *et al.*, 2019; Roca-Barcelo *et al.*, 2020). Studies on occupational exposure to bacteria, fungi, and fine particles, particularly those with a diameter of less than 10 µm, have shown that these factors have an impact on people's health. (Mack *et al.*, 2019). Individuals exposed to organic dust from composting facilities were more likely to experience irritation of the mucosal membranes in their eyes

and upper airways (Bunger *et al.*, 2007). Also, prolonged exposure to bioaerosols is associated with acquired allergic lung disease, as reported by Flannigan *et al.* (Flannigan *et al.*, 1991). In England, studies by Douglas *et al.* and Roca-Barcelo *et al.* found no evidence of a higher incidence of respiratory admission among the general population living close to large composting sites (Douglas *et al.*, 2016a; Roca-Barcelo *et al.*, 2020).

Living close to major composting sites might have a detrimental impact on individuals with respiratory conditions like asthma, COPD and CF in particular, and exposure to bioaerosols may provide a significant health concern for these high-risk pwCF (Walser *et al.*, 2015). In pwCF, thick secretions form in the lungs, creating an ideal condition for the development of bacteria and fungus infections. As a result, these pwCF may be more vulnerable to frequent PEx and respiratory infections (Ranganathan *et al.*, 2013). Respiratory infections are one of the leading causes of morbidity and mortality in the CF population. Significant research has been done on pwCF's bacterial infections, including their treatment and prevention (Chiappini *et al.*, 2014; Gilligan, 2014). While this has helped increase life expectancy, it may also have increased the likelihood of fungus colonisation in the CF population (Gilligan, 2014). Fungi are the most widespread species, and among pwCF, *A. fumigatus* is the most prevalent fungus (Burgel *et al.*, 2016). The prevalence of *A. fumigatus*-positive cultures increased from 6% to 15% between 2007 and 2016, according to data from the UK CF Registry (UK CF Registry., 2020). There is evidence that various fungal infections are correlated with a drop in lung function and a higher number of PEx (Saunders *et al.*, 2016).

Although the underlying process is still unclear, environmental influences are expected to play a significant role (Crawford *et al.*, 2015; Mendell *et al.*, 2011). Additionally, it's still not understood how the bioaerosols affect the health of pwCF, who are particularly vulnerable. It is important to note that no quantitative dose-response estimation is currently available to inform laws about bioaerosol emissions from PCS. There is limited evidence from the UK on the effect of composting sites on PEx and lung function in pwCF. FEV_1 and PEx are used to monitor the progress of the disease and detect any problem that deteriorates the pwCF's health status. Understanding the frequency of infections in pwCF living near composting facilities may help in identifying environment-related risk factors and how they affect health (Robertson *et al.*, 2019). To the best of my knowledge, no research has looked at how living near large composting sites may affect lung function and PEx in pwCF in the UK. This research

will help better understand the influence of where individuals with CF live and the disease severity of pwCF.

4.1.1 Hypothesis

I hypothesised that the bioaerosols emitted by PCSs are associated with a lower lung function, an increase in PEx, and increased bacterial and fungal infection in pwCF in the UK.

4.1.2 The objective of the study

This national-scale study examined the potential associations between distance to PCS, a bioaerosol exposure proxy, and changes in lung function, PEx, bacterial and fungal infections in pwCF.

4.2 Materials and methods

4.2.1 Study design, population and the outcome variable

The study's design, population, exclusion criteria, and outcome variables are all thoroughly covered in ***Chapter II, Section 2.1.1 – 2.2.10***. I will briefly summarise here. The effect of living close to a PCS on lung function, PEx, and bacterial and fungal illness in children and adults with CF was examined using a semi-individual cross-sectional study design. I included all pwCF who had an annual review in 2016 in this research study using data from the UK CF Registry. I considered three key outcome variables: lung function, PEx, and bacterial and fungal disease. The first two outcome variables are similar to those used in the Chapter II analysis. I chose ppFEV$_1$ and PEx as my outcome measures for several reasons. Firstly, they directly reflect the impact of environmental exposures on lung function and disease severity in CF (Schechter, 2011; Szczesniak *et al.*, 2020). Secondly, their significance is supported by existing research and discussion with clinical and field experts: Professor Nicholas Simmonds and Dr Philippa Douglas. Importantly, the availability of relevant data in the CF registry allows for robust analysis. These outcomes are plausibly linked to composting exposure through potential inhalation of pathogens like *P. aeruginosa and S. aureus*, as well as the inflammatory effects of bioaerosols themselves (Breza-Boruta, 2016; Wéry, 2014). While individual susceptibility may play a role, investigating these measures within the CF registry offers a powerful opportunity to understand the potential risks of composting site proximity on the health of pwCF.

Another significant outcome variable that was taken into account in this analysis was bacterial and fungal disease. I used positive fungal cultures (sputum, throat/cough swab, or bronchoalveolar lavage) of *A. fumigatus* and *Candida spp.* in the year prior to the yearly review and/or a record of ABPA to determine the existence of a fungal infection or disease. The UK CF Registry defines ABPA in accordance with accepted global diagnostic standards, which include an *Aspergillus* antigen skin prick test result of greater than 3 or a positive result from a specific IgE test for *A fumigatus*, as well as acute or subacute clinical symptoms (cough, wheeze, exercise intolerance, exercise-induced asthma, decrease in lung function, or increased sputum production) not due to another aetiology. Additionally, either precipitating *A. fumigatus*, with in vitro IgG antibody testing for *A. fumigatus*, or any new or

recent abnormalities on chest radiograph (infiltration or mucus plugging) or CT scan (characteristic changes) that have not resolved with treatment with antibiotics and conventional physiotherapy (Patel *et al.*, 2019; Stevens *et al.*, 2003). Additionally, I looked at *P. aeruginosa* and *S. aureus* infections that were reported in the UK CF Registry and had a culture-positive result in the year before the yearly review.

4.2.2 Exposure variable

A permit is required to operate a commercial composting facility in the UK. According to DEFRA, a permit is needed in England if a facility handles more than 60,000 kilogrammes of compost at one time; however, different weight limits are used in Wales, Scotland, and Northern Ireland (DEFRA., 2014). To obtain comprehensive data on all composting sites in the UK, I contacted all four Environment Agencies operating within the UK, namely the UK Environment Agency (England), Natural Resources Wales, the Scottish Environment Protection Agency, and the Northern Ireland Environment Agency, as these are the main regulatory authorities and responsible for monitoring and managing environmental pollution in their respective areas. All functioning PCSs in the UK that were functioning as of the end of 2016 were selected for analysis for this research study.

The site postcode provided by environment agencies was used to geocode the PCS in ArcGIS, which converts the postcode into geographic latitude and longitude for an accurate location of the PCS on the map (ESRI., 2018). ***Figure 2*** displays the geographical distribution of PCSs around the UK. In ArcGIS, the distance between the nearest PCS and the pwCF's residence postcode was calculated using the Euclidian distance (point distance analysis).

I explored various circular distance bands (0 - 250 m, >250 - 750 m, >750 - 1.5 km, >1.5 - 2.5 km, >2.5 - 4.0 km, and > 4.0 km) surrounding each PCS by earlier studies; however, I focused this investigation within and beyond 4 kilometres of the PCS due to the small number of pwCF living in the distance bands adjacent to the PCS (Douglas *et al.*, 2016a; Robertson *et al.*, 2019; Roca-Barcelo *et al.*, 2020).

4.2.3 Covariates

Age, sex, *CFTR* genotype, BMI, *P. aeruginosa*, urban-rural categorization, and socioeconomic status level (SES) were the variables identified a *priori* as potential confounders. In ***Chapter II, section 2.2.10***, each of these has been described in depth.

4.2.4 Statistical analysis

Statistical and spatial analyses were conducted using *R* 4.2.2 (R Core Team., 2022) and ArcGIS 10.7.1 (ESRI., 2018), respectively.

To explore the distribution of the data, a descriptive analysis of variables (demographic and clinical characteristics) was performed. Two-way ANOVAs were used to examine the descriptive characteristics of pwCF living within and beyond 4 km of a PCS, including age, sex, *CFTR* genotype, and *P. aeruginosa*, as well as area-level descriptors including deprivation (Townsend index), rural-urban categorization, and country. Regression models for adults and children were created separately. The relationship between $ppFEV_1$ and distance (≤ 4 km versus > 4 km) from a PCS was examined using a multiple linear regression analysis while controlling for age, sex, genotype, BMI*, P. aeruginosa*, and deprivation, With the use of Poisson regression, I was able to explore the association between the #IVdays for PEx and the distance from a PCS. The models were selected in a forward-stepwise manner, with covariates being added one at a time using the Akaike information criterion (AIC) (Burnham & Anderson, 2003). I chose the model with the lowest AIC. (***Table 4.4***).

Infection rates for *A. fumigatus*, ABPA, *Candida spp.*, *P. aeruginosa*, and *S. aureus* for pwCF living within and beyond 4 km of a PCS were computed and reported. An ANCOVA was used to determine the adjusted mean $ppFEV_1$ for each of these five indicators of both bacterial and fungal infection in CF children as well as CF adults. The threshold for statistical significance was set at a p-value of 0.05; 95% confidence intervals (CIs) and p-values are presented.

4.2.5 Ethics and approval

This is explained in ***Section 2.2.12*** of Chapter II.

4.3 Results

In 2016, annual review data was available for 9,695 pwCF in total. A schematic summary of eligible pwCF for this study is shown in ***Figure 4.1***. For the investigation of PEx, bacterial and fungal diseases, data on 9,631 pwCF were available, comprising 3,931 CF children and 5,430 CF adults. There were 7,800 pwCF who remained for the $ppFEV_1$ analysis after removing those under the age of six. ***Figure 4.2*** shows that in 2016, there were 256 active PCSs in the UK, with 209 of them in England, 24 of them in Wales, 21 of them in Scotland, and 2 of them in Northern Ireland. ***Table 4.1*** summarises the number of pwCF within various distance bands. Analyses were performed for these various distances. However, the statistical power of analysis in smaller bands was ultimately limited by the few pwCF residing close to PCS, therefore, I narrowed the scope of my investigation to areas within and beyond 4 km of PCS; 1,015 pwCF (10.8%) had a residential postcode within 4 kilometres of a PCS.

Tables 4.2 and 4.3 summarise the demographics of pwCF living within and beyond 4 km of a PCS. pwCF had a mean age of 20.1 (SD: 14.1) years; 49% (n = 4,588) of them were between the ages of 16 and 40; 53.3% of them were male; and 19.2% of them lived in rural areas. There was a nearly equal distribution of pwCF in all the quintiles of deprivation. England had a pwCF distribution of 7,675 (82.0%), Wales had 494 (5.3%), Scotland had 796 (8.5%), and Northern Ireland had 396 (4.2%). According to *CFTR* genotypes, 6.2% of pwCF had a gating mutation, 11.4% had an RF mutation, 24.3% had at least one MF mutation, and 49.2% were homozygous for the *F508del* mutation.

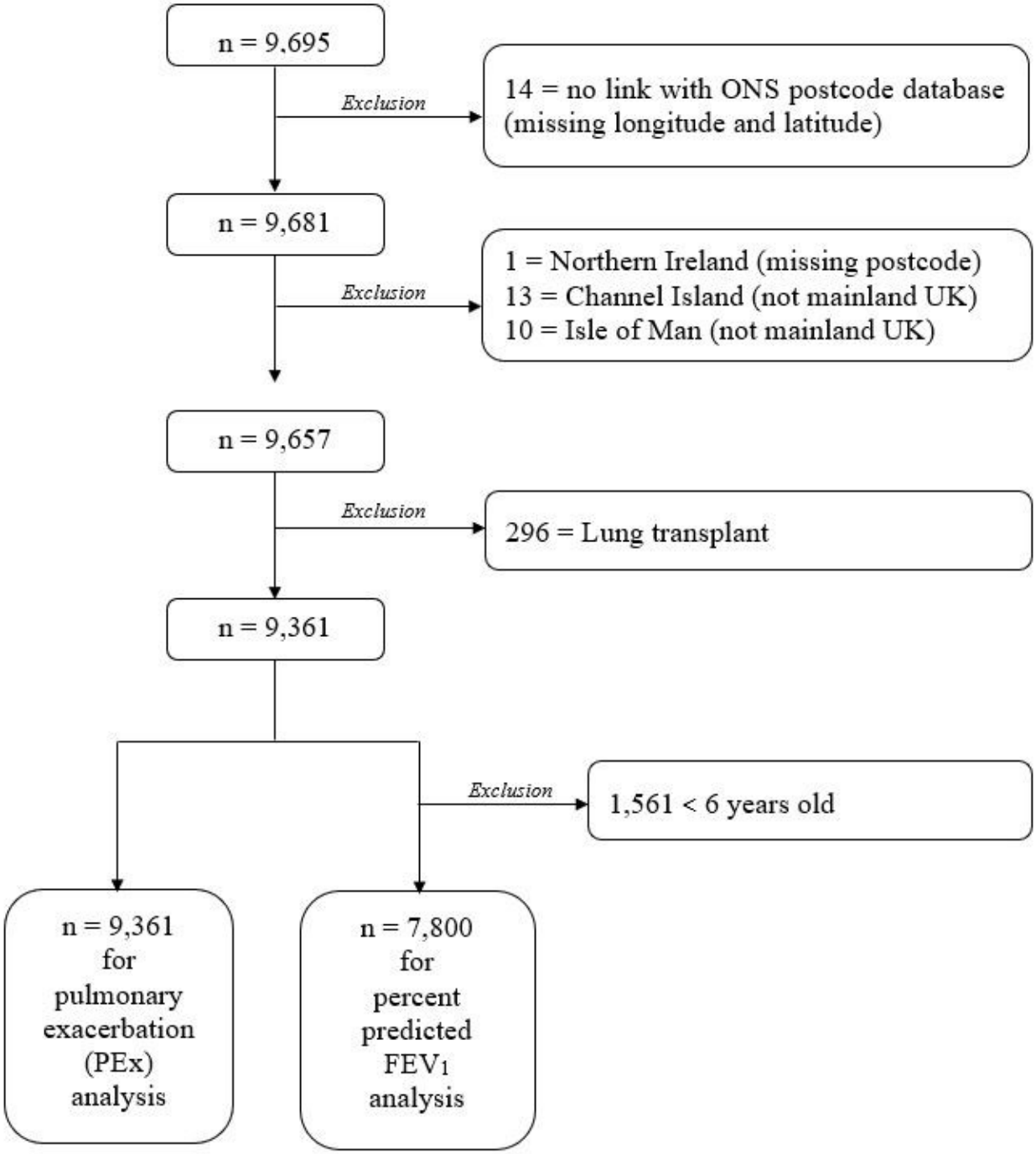


Figure 4.1: Flow diagram of the inclusion and exclusion criteria used for people with CF (pwCF) in the 2016 UK CF Registry database

Table 4.1: Number of people with CF (pwCF) living within specific distances from a permitted composting site (PCS) in the UK in 2016

Distance to a PCS	***pwCF***	***%***
0-250 m	0	0.0
> 250 - ≤ 750 m	18	0.2
> 750 - ≤ 1.5 km	89	1.0
> 1,5 - ≤ 2.5 km	260	2.8
> 2.5 - ≤ 4.0 km	649	6.9
> 4.0 km	8,345	89.1
TOTAL	9,361	100.0

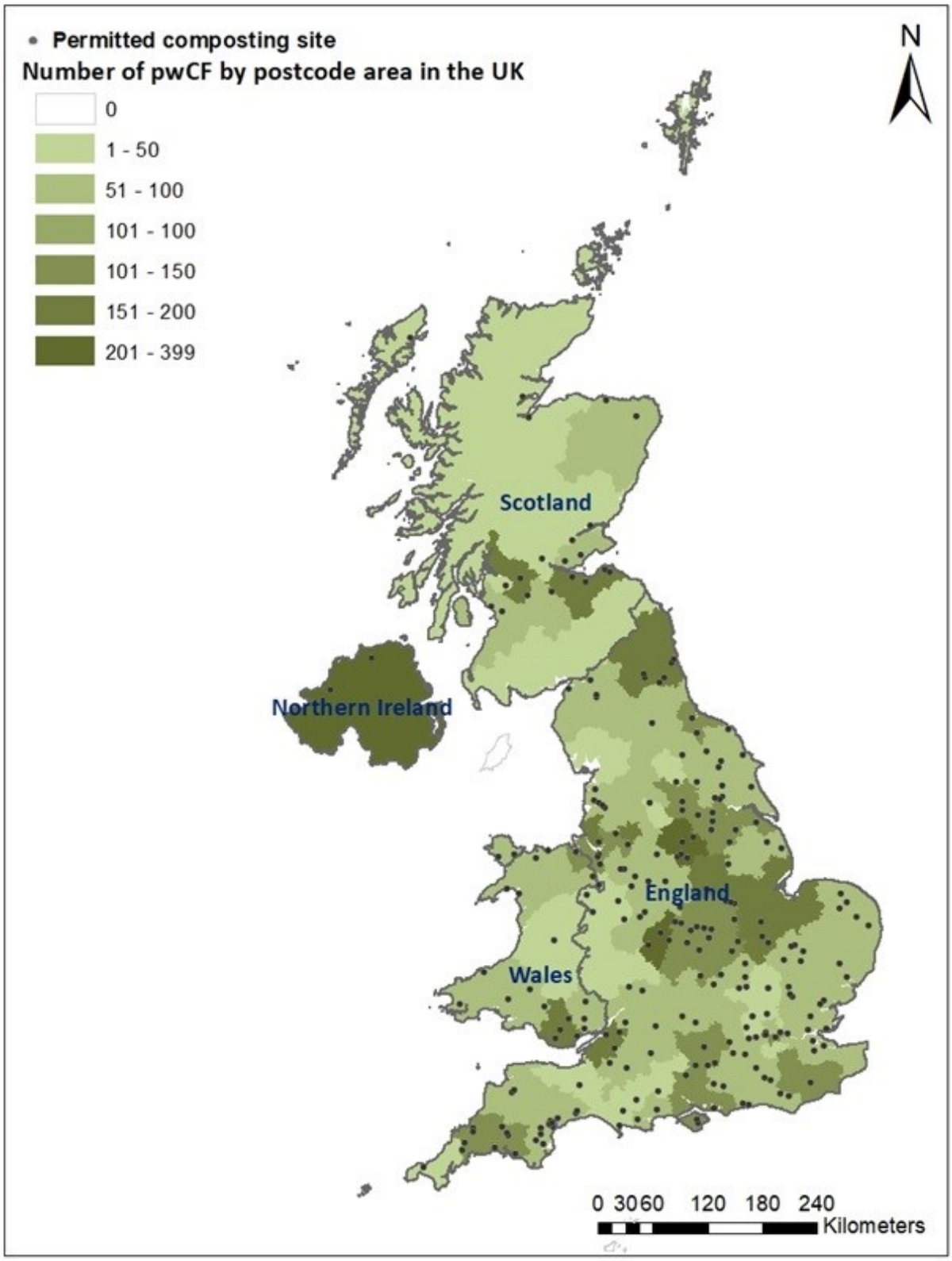


Figure 4.2: Map of the number of pwCF by postcode area and permitted composting sites (PCSs), in the UK in 2016 based on the centroid of postcodes provided by the Environment Agencies of England (n=209), Scotland (n=21), Wales (n=24) and Northern Ireland (n=02)

Table 4.2: Mean $ppFEV_1$ within and beyond 4 km of a PCS in pwCF in the UK in 2016
p-value <0.05 are shown in bold.

Characteristics	n (%)	Mean $ppFEV_1$ mean (*SD*) ≤ 4 km	mean (*SD*) > 4 km	*p*-value
	n=7,800 (100%)	(n=866; 11.1 %)	(n= 6,934; 88.9%)	
Age (years)				
Children < 6	-*			**0.0265**
6 - 15	2,370 (30.4)	85.0 (16.7)	86.1 (17.2)	
Adults ≥ 16	5,430 (69.4)	62.2 (23.7)	64.3 (23.6)	
Sex				
Female	3,629 (46.5)	66.7 (24.5)	70.8 (23.8)	**<0.001**
Male	4,171 (53.5)	69.1 (24.0)	71.2 (24.3)	
Genotype				
homozygous *F508del*	3,825 (49.0)	66.1 (24.5)	68.9 (24.0)	
F508del/MF	1,511 (19.4)	66.3 (23.2)	69.0 (24.1)	
RF/other	874 (11.2)	75.9 (24.6)	79.1 (22.5)	
F508del/other	558 (7.2)	70.3 (23.1)	73.7 (24.6)	**0.001**
Gating/any**	494 (6.3)	73.3 (25.4)	75.9 (23.5)	
MF/MF	249 (3.2)	67.0 (18.5)	69.1 (21.9)	
Other/other	158 (2.0)	77.5 (21.9)	73.3 (23.6)	
MF/other	131 (1.7)	63.1 (28.3)	76.4 (22.8)	
***P. aeruginosa* culture**				
Negative	3,984 (51.1)	76.3 (22.3)	79.5 (21.7)	**0.0016**
Positive	3816 (48.9)	60.4 (23.5)	62.4 (23.3)	
Deprivation quintile (Townsend index)				
Q1 (least deprived)	1,612 (20.7)	67.8 (23.6)	73.3 (23.6)	
Q2	1,615 (20.7)	65.8 (26.1)	71.1 (24.7)	**0.002**
Q3	1,666 (21.4)	68.7 (22.9)	70.8 (24.1)	
Q4	1,547 (19.8)	69.2 (24.3)	70.3 (23.5)	
Q5 (most deprived)	1,360 (17.4)	68.3 (24.6)	69.3 (24.1)	
Rural and Urban				
Rural	1,500 (19.2)	66.2 (24.6)	72.4 (24.1)	**0.001**
Urban	6,300 (80.8)	68.3 (24.2)	70.7 (24.0)	

$ppFEV_1$ = percent predicted FEV_1; PCS = permitted composting site; pwCF = people with cystic fibrosis; *Pseudomonas aeruginosa = P. aeruginosa;* *FEV_1 not reliable under six years of age hence excluded from $ppFEV_1$ analysis; MF = Minimal function (a mutation that produces either no protein or protein that do not respond to currently approved *CFTR* modulators); RF = Residual function; Other = Either the mutation was not identified, or it was identified but classification into one of the other categories were not possible. **Except RF mutation

Table 4.3: Mean #IVdays within and beyond 4 km of a PCS in pwCF in the UK in 2016
p-value <0.05 are shown in bold.

Characteristics	#IVdays			
	n (%)	**mean (*SD*)**	**mean (*SD*)**	
		≤ 4 km	**> 4 km**	***p*-value**
	(n=9,361 (100%)	(n=1,015;10.8%)	(n=8,346; 89.2%)	
Age (years)				
Children < 6	1,561 (16.7)	5.5 (9.8)	5.1 (11.2)	0.291
6 - 15	2,370 (25.3)	12.7 (21.8)	12.4 (20.9)	
Adults ≥ 16	5,430 (58.0)	22.4 (34.2)	21.0 (32.7)	
Sex				
Female	4,371 (46.7)	22.0 (35.1)	18.7 (30.5)	**0.003**
Male	4,990 (53.3)	14.2 (24.4)	13.9 (25.4)	
Genotype				
homozygous *F508del*	4,603 (49.2)	22.9 (38.1)	17.7 (29.9)	
F508del/MF	1,822 (19.5)	12.3 (21.7)	12.4 (25.0)	
RF/other	1,070 (11.4)	8.4 (22.1)	8.4 (17.0)	
F508del/other	663 (7.1)	21.0 (30.1)	19.4 (30.1)	**0.004**
Gating/any*	583 (6.2)	11.8 (19.6)	19.0 (27.6)	
MF/MF	298 (3.2)	18.7 (30.3)	11.0 (20.2)	
Other/other	176 (1.9)	10.6 (22.8)	12.2 (32.3)	
MF/other	146 (1.6)	5.7 (19.7)	6.2 (16.5)	
***P. aeruginosa* culture**				
Negative	5,108 (55.2)	77.6 (22.4)	80.4 (21.4)	**<0.001**
Positive	4,139 (44.8)	61.2 (24.0)	62.8 (23.4)	
Deprivation quintile (Townsend index)				
Q1 (least deprived)	1,905 (20.4)	15.2 (24.6)	14.1 (25.4)	
Q2	1,948 (20.8)	16.8 (27.1)	15.6 (28.1)	**0.005**
Q3	1,982 (21.2)	15.7 (32.1)	15.4 (26.4)	
Q4	1,868 (20.0)	17.7 (30.5)	16.6 (28.8)	
Q5 (most deprived)	1,658 (17.7)	22.0 (33.6)	19.4 (31.5)	
Rural and Urban				
Rural	1,794 (19.2)	16.0 (25.1)	15.3 (26.4)	**0.004**
Urban	7,567 (80.8)	18.1 (30.7)	16.3 (28.4)	

#IVdays = Number of days on IV antibiotics for PEx; PCS = permitted composting site; pwCF = people with cystic fibrosis; *Pseudomonas aeruginosa = P. aeruginosa;* MF = Minimal function (a mutation that produces either no protein or protein that do not respond to currently approved *CFTR* modulators); RF = Residual function; Other = Either the mutation was not identified, or it was identified but classification into one of the other categories were not possible. *Except RF mutation

4.3.1 Lung function

The mean $ppFEV_1$ of the paediatric population was 62.2% within 4 km of a PCS and 64.3% beyond 4 km (p = 0.044); whereas the adult population had a mean $ppFEV_1$ of 85.0% within 4 km of a PCS and 86.1% beyond 4 km (p = 0.391). The pwCF living within 4 km of a PCS were significantly different from the pwCF living farther away in terms of age (p = 0.027), sex (p < 0.001), *CFTR* genotype (p = 0.001), deprivation index (p = 0.002) and rural proportion (p = 0.001) ***(Table 4.2).*** The mean $ppFEV_1$ of CF children aged 6 to 15 living within and beyond a PCS's 4-kilometre boundary did not show any statistically significant association in this investigation (p = 0.358) ***(Table 4.4)*****.** In my univariate linear regression analysis, I found that CF adults who lived within 4 km of a PCS had a lower $ppFEV_1$ than the rest of the CF adults in the study by -1.44% (95% CI: -2.84; -0.04). When age, sex, genotype, BMI, *P. aeruginosa* and deprivation were taken into account in the multiple linear regression model, the difference in $ppFEV_1$ became smaller, and the association was weakened by -1.07% (95% CI: -2.29; 0.16, p = 0.088). ***Table 4.5*** summarises various models for the association between $ppFEV_1$ and proximity to a PCS based on AIC.

Table 4.4: Linear and Poisson regression analysis of the relationship between ppFEV$_1$ and #IVdays to residential postcode distance from PC in 2016.

p-value <0.05 are shown in bold.

Residential postcode distance	ppFEV$_1$					#IVdays					
	mean ppFEV$_1$	**Coeff.**	**95% CI**	***p-value***	**AIC**	**mean #IVdays**	**Coeff.**	**Exp (B)**	**95% CI**	***p-value***	**AIC**
Children (6 -15 years) (n=2,370)						**Children (0 -15 years) (n=3,931)**					
Unadjusted				0.391	19297.9					0.087	103952.3
≤ 4 km	85.0	-0.75	-2.46, 0.96			9.8	0.02	1.02	0.99, 1.05		
> 4 km	86.1	1.00	-			9.5		1.00	-		
Adjusted*				0.358	18887.6					0.623	84834.0
≤ 4 km	86.0**	-0.74	-2.31, 0.83			10.3	0.01	1.01**	0.99, 1.04		
> 4 km	86.8**	1.00	-			10.3		1.00**			
Adults (16 - 86 years) (n=5,430)						**Adults (16 - 86 years) (n=5,430)**					
Unadjusted				**0.044**	47438.5					**<0.001**	231519.0
≤ 4 km	62.2	-1.44	-2.84, -0.04			22.4	0.04	1.04	1.03, 1.06		
> 4 km	64.3	1.00	-			21.0		1.00	-		
Adjusted*				0.088	46040.1					**<0.001**	187492.7
≤ 4 km	64.4**	-1.07	-2.29, 0.16			19.4	0.02	1.02**	1.01, 1.04		
> 4 km	66.3**	1.00	-			18.4		1.00**			

ppFEV$_1$ = percent predicted FEV$_1$; #IVdays = Number of days on IV antibiotics for PEx; pwCF = people with cystic fibrosis; PCS = permitted composting site

*Adjusted for age, sex, BMI (children model adjusted for BMI percentile and adult model adjusted for absolute BMI), genotype, *P. aeruginosa* and deprivation

** calculated using ANCOVA (Analysis of Covariance) adjusting for age, sex, BMI (children model adjusted for BMI percentile and adult model adjusted for absolute BMI), genotype, *P. aeruginosa* and deprivation

Table 4.5: Linear regression analysis of the relationship between ppFEV$_1$ and residential postcode distance from PCS among children and adults with CF in the UK in 2016
p-value <0.05 are shown in bold.

	Children (6 -15 years) (n=2,370)		**Adult (16 - 86 years)** (n=5,430)	
ppFEV$_1$	≤ 4 km (n=222)	> 4 km (n=2148)	≤ 4 km (n=644)	> 4 km (n=4786)
Unadjusted				
β (95%CI)	-0.75 (-2.45, 0.96)	Ref	-1.44 (-2.84, -0.04)	Ref
p-value	0.391		**0.044**	
AIC	19297.9		47438.5	
Adjusted model I				
β (95%CI)	-0.90 (-2.54, 0.74)	Ref	-1.47 (-2.80, -0.14)	Ref
p-value	0.282		**0.031**	
AIC	19118.11		46929.3	
Adjusted model II				
β (95%CI)	-1.11 (-2.69, 0.47)	Ref	-1.50 (-2.78, -0.22)	Ref
p-value	0.168		**0.022**	
AIC	18935.0		46535.8	
Adjusted model III				
β (95%CI)	-1.06 (-2.63, 0.51)	Ref	-1.34 (-2.60, -0.10)	Ref
p-value	0.184		**0.034**	
AIC	18910.4		46185.4	
Adjusted model IV				
β (95%CI)	-1.00 (-2.56, 0.57)	Ref	-1.20 (-2.42, 0.03)	Ref
p-value	0.211		0.056	
AIC	18896.5		46062.6	
Adjusted model V				
β (95%CI)	-0.71 (-2.8, 0.87)	Ref	-1.05 (-2.28, 0.17)	Ref
p-value	0.377		0.092	
AIC	18887.7		46041.3	
Adjusted model VI				
β (95%CI)	-0.74 (-2.31, 0.83)	Ref	-1.07 (-2.29, 0.16)	Ref
p-value	0.343		0.088	
AIC	18887.6		46040.6	

ppFEV$_1$ = percent predicted FEV$_1$; CF = Cystic fibrosis; PCS = permitted composting site

Children
Model I adjusted for BMI percentile
Model II adjusted for BMI percentile and age,
Model III adjusted for BMI percentile, age and genotype
Model IV adjusted for BMI percentile, age, genotype and *P. aeruginosa* sputum culture-positive
Model V adjusted for BMI percentile, age, genotype and *P. aeruginosa* sputum culture positive and deprivation
Model VI adjusted for BMI percentile, age, genotype and *P. aeruginosa* sputum culture positive, deprivation and sex

Adults
Model I adjusted for *P. aeruginosa* sputum culture-positive
Model II adjusted for *P. aeruginosa* sputum culture positive and absolute BMI
Model III adjusted for *P. aeruginosa* sputum culture positive, absolute BMI and age
Model IV adjusted for *P. aeruginosa* sputum culture positive, absolute BMI, age and genotype
Model V adjusted for *P. aeruginosa* sputum culture positive, absolute BMI, age, genotype and deprivation
Model VI adjusted for *P. aeruginosa* sputum culture positive, absolute BMI, age, genotype, deprivation and sex

4.3.2 PEx

Children who lived within and beyond 4 km of a PCS had an average of 9.8 and 9.5 days of # IV antibiotics each year, respectively, whereas adults had an average of 22.4 and 21.0 days of IV each year. In univariate analysis, I found a significant difference in the #IVdays between sex (p = 0.003), genotype (p = 0.004), *P. aeruginosa* (p < 0.001), deprivation quintiles (p = 0.005), and rural and urban locations (p = 0.004) for pwCF living within and beyond 4 km of the PCS (***Table 4.3***). In my Poisson regression analysis, I did not find a significant association between the #IVdays among CF children and the distance to PCS. In CF adults who lived within and beyond 4 km of a PCS, I found that the rate of #IVdays per year rose by 1.04 (95%CI: 1.03; 1.06) days and 1.02 (95%CI: 1.01; 1.04) days, respectively, before and after accounting for age, sex, BMI, genotype, *P. aeruginosa* sputum culture positive and deprivation (***Table 4.3***).

4.3.3 Fungal and bacterial infections

In ***Table 4.6*** I demonstrate that there was no significant difference in the prevalence of culture-positive *A. fumigatus* among all pwCF living both within and beyond 4 km of a PCS (15.3% *vs.* 15.5%, respectively). However, the mean FEV_1 of adults with CF who have a positive *A. fumigatus* culture and live within 4 km of PCS was significantly lower than that of those who lived farther away (55.9% *vs.* 62.1%: p = 0.001). This difference was still significant when age, sex, genotype category, BMI and deprivation were taken into account (58.2% *vs.* 62.0%; p = 0.005). Additionally, *Candida spp.* point prevalence was also higher among those living close to a PCS (3.6%, p = 0.01). In the CF adults with a positive culture for *Candida spp.,* the difference in mean $ppFEV_1$ remained significant (56.9% *vs*. 59.0%; p = 0.029) after accounting for age, sex, absolute BMI, genotype, and deprivation. Adults with CF living within and beyond 4 km of a PCS did not have a significantly different point prevalence of ABPA, *P. aeruginosa*, or *S. aureus*. Additionally, I did not find a statistically significant difference in the mean $ppFEV_1$ of CF children or CF adults who had ABPA or a positive culture of *P. aeruginosa* or *S. aureus* and who lived within and beyond 4 km of a PCS (***Table 4.6***).

Table 4.6: Respiratory microbiology, fungal disease and mean ppFEV$_1$ among children and adults with CF within and beyond 4km of PCS in the UK in 2016
p-value <0.05 are shown in bold.

	Distance band Crude			Distance band Adjusted*		
	≤ 4 km	**> 4 km**	***p-value***	**≤ 4 km**	**> 4 km**	***p-value***
Total Number of pwCF (n)	866	6,934	-	-	-	-
Number of pwCF who had culture* (n)	842	6,730	-	-	-	-
A. fumigatus						
n (point prevalence %)	129 (15.3)	1,041 (15.5)	0.911			
mean ppFEV$_1$						
Children	77.2	79.0	0.662	78.3	80.4	0.631
Adults	55.9	62.1	**0.001**	58.2	62.0	**0.005**
ABPA						
n (point prevalence %)	76 (8.8)	551 (7.9)	0.397			
mean ppFEV$_1$						
Children	81.2	75.3	0.229	86.6	82.6	0.179
Adults	60.2	59.1	0.717	60.3	60.1	0.696
Candida spp.						
n (point prevalence %)	180 (21.4)	1,195 (17.8)	**0.010**			
mean ppFEV$_1$						
Children	79.6	79.3	0.929	84.0	82.9	0.921
Adults	53.6	57.8	**0.048**	56.9	59.9	**0.029**
P. aeruginosa						
n (point prevalence %)	439 (52.1)	3,377 (50.2)	0.284			
mean ppFEV$_1$						
Children	80.5	82.3	0.436	80.9	84.1	0.399
Adults	56.8	58.1	0.280	58.5	59.6	0.240
S. aureus						
n (point prevalence %)	76 (8.3)	558 (9.0)	0.469			
mean ppFEV$_1$						
Children	71.2	76.1	0.541	63.7	70.8	0.472
Adults	58.0	56.9	0.690	59.1	59.0	0.663

Aspergillus fumigatus = A. fumigatus; Pseudomonas aeruginosa = P. aeruginosa; Staphylococcus aureus = S. aureus
*(Sputum/ Cough Swab/ Bronchoscopy); ppFEV$_1$ = percent predicted with FEV1; pwCF = people with cystic fibrosis; ABPA = Allergic Bronchopulmonary Aspergillosis; PCS = permitted composting site; Children = 6 - 15 years; adult = 16 - 86 years
*Adjusted for age, sex, BMI (children model adjusted for BMI percentile and adult model adjusted for absolute BMI), genotype and deprivation

4.4 Discussion

This is the first nationwide research that has investigated the potential impact of living close to a PCS for pwCF, in terms of their lung function, PEx, and risk of bacterial and fungal infections. Additionally, this study investigated the prevalence of positive fungal and bacterial cultures within and beyond a 4-kilometre distance band around a PCS. These results demonstrated a relatively lower ppFEV$_1$ (1.07%) and higher #IVdays (1.02 days/year) in adults with CF who lived near a PCS (≤ 4 km); however, there was no significant association between the disease severity and children's proximity to a PCS. Given that low FEV$_1$ is strongly associated with both high mortality and poor quality of life in CF, it is important to understand the magnitude of the FEV$_1$ decrease that is considered clinically significant and relevant. Due to its reliability and reproducibility, FEV$_1$ has been used as the key outcome variable to monitor pwCF's progress in many CF clinical trials (Ramsey *et al.*, 2011; Wainwright *et al.*, 2015). For example, in a phase 3 randomised double-blind trial, Wainwright *et al.* found a statistically significant improvement in ppFEV$_1$ in the range of 2.6% and 4.0% in a clinical trial evaluating the efficacy of lumacaftor and ivacaftor (VX-809) among 1,108 pwCF ($p < 0.001$) (Wainwright *et al.*, 2015). PEx are strongly associated with a significant decline in FEV$_1$ and a poor quality of life in pwCF. Additionally, they are also thought to be a significant cause of morbidity. Even though a small drop in the mean ppFEV$_1$ and a one-day drop in #IVdays might not be clinically significant in an individual, they might indicate a major and significant difference at a population level.

The distance of 4 km was selected in accordance with the distance bands used in earlier research that were reported in peer-reviewed publications that used distance as a proxy for exposure to bioaerosol near composting sites and its relationship to health effects (Roca-Barcelo *et al.*, 2020; Williams *et al.*, 2019). There are few community exposure studies that look at the bioaerosols that come from composting sites, and while these studies do indicate that bioaerosols drop to background levels within 2.5 km of a composting site, they have many limitations, which are detailed in two systematic reviews (Pearson *et al.*, 2015; Robertson *et al.*, 2019).

There has not been much prior research on how living near PCSs affects people's health in general or individuals affected by respiratory diseases such as CF. However, there are

occupational and community studies that explore the effect of bioaerosol on the respiratory system. According to research by Bünger *et al.,* German composting employees were more likely to have chronic bronchitis, mucosal membrane irritation in the eyes and upper airways, and their Forced Vital Capacity had significantly decreased (Bunger *et al.*, 2007). Similarly, Hambach *et al.* found more respiratory, skin and gastrointestinal symptoms among composting workers than a control group with a comparable socio-economic status in Belgium (Hambach *et al.*, 2012). Other occupational research studies also reported a strong association between composting workers and respiratory symptoms (Hoffmeyer *et al.*, 2014; Müller *et al.*, 2006; van Kampen *et al.*, 2016). A comprehensive meta-analysis of various occupational and community research studies that looked at bioaerosols' effect on respiratory health has presented these findings (Pearson *et al.*, 2015). Additionally, other studies also found a strong association between lung function and organic dust (Jacobsen *et al.*, 2008; Senthilselvan *et al.*, 2010). Bolund AC *et al.* reported a significant decrease of 4.92 mL/year (95% CI: 0.14; 9.69) in FEV_1 among those exposed to organic dust compared to controls (Bolund *et al.*, 2017).

There have been some investigations into the relationship between respiratory health and proximity to PCS. In Germany, Herr *et al.* reported that individuals who lived within 150 – 200 metres of a PCS had significantly higher odds of having bronchitis (OR = 3.59, 95%CI: 1.40; 9.47), coughing during the day (OR = 3.18, 95%CI: 1.24; 8.36) and waking up due to coughing (OR = 6.59, 95%CI: 2.57; 17.7) compared to those who lived more than 400–500 metres away (Herr *et al.*, 2003). Similarly, Aatamila *et al.* also observed unusual breathing difficulty (OR = 1.5, 95% CI: 1.0; 2.2) and unusual fatigue (OR = 1.5, 95% CI: 1.1; 2.0) among those living within distance bands of less than 1.5 km and individuals living between 3.0 and 5.0 km from waste treatment centres in Finland (Aatamila *et al.*, 2011). This could be due to using hospital admission as the key outcome variable, which has an extreme spectrum of health risks from exposure to bioaerosol. Individuals with initial respiratory symptoms might have been sorted at the primary care level rather than the hospital level. Nevertheless, in England, Douglas *et al.* did not observe a higher risk of respiratory hospital admissions in the general population living close to a PCS, using distance from a PCS as a proxy for exposure to bioaerosol (Douglas *et al.*, 2016a).

Furthermore, a subsequent investigation in England by Roca-Barcelo *et al.* did not observe any respiratory-related health effects or a higher risk of CF hospitalisation in the areas with

higher *A. fumigatus* exposure (calculated using the dispersion model) close to PCSs (Roca-Barcelo *et al.*, 2020). Environmental factors may have a substantial impact on the onset of PEx, with serious implications and effects on pwCF's health. The pathophysiology of pulmonary disease in pwCF may be better understood if the environmental risk factors and negative effects, especially the way they mediate them and affect disease development, are better understood.

4.4.1 Fungal infection

The mean ppFEV$_1$ was significantly lower among pwCF with a positive fungal culture living close to a PCS. In comparison to CF individuals living further away from a PCS, those with positive *A. fumigatus* cultures had a mean ppFEV$_1$ that was 5.2% lower, and those with positive *Candida spp.* cultures had a mean ppFEV$_1$ that was 4.2% lower. This research is unable to provide information about the source of the pathogens, a probable explanation for this, or even whether it is related to unmeasured factors. One possibility can be a higher burden of fungal load (density) in the sputum of pwCF living near PCSs (Williams *et al.*, 2019). According to Williams *et al.,* high anthropogenic *A. fumigatus* model loads were found within a 3 to 4 km area of a PCS in England (Williams *et al.*, 2019). Children who lived close to a PCS had a higher prevalence of ABPA, albeit this association was not statistically significant, which may offer some indirect support for this argument since a larger concentration of fungus spores in the air could increase sensitization. This must be tested in a robust prospective study. Fungal burden in sputum could be determined using quantifiable molecular microbiological approaches to detect fungi at the level of species (Cuthbertson *et al.*, 2021). Most research to date has only used conventional culture-based techniques to examine the relationship between lung function and cultures positive for fungi in sputum (Burgel *et al.*, 2016; Kraemer *et al.*, 2006). For example, Amin *et al.* found that the mean FEV$_1$ values for pwCF with positive and negative cultures of Aspergillus were 79.2% and 86.1% (p = 0.04), respectively (Burgel *et al.*, 2016). Similar findings were reported by Kraemer *et al.* who found that CF children with positive cultures of *Aspergillus* had significantly lower lung function than CF controls (Kraemer *et al.*, 2006).

4.4.2 Strengths and limitations of the study

The key strength of this research is that I used high-quality UK CF Registry data, which is believed to capture over 99% of pwCF in the UK and included all the UK's active composting sites in 2016. The UK CF Registry data contains in-depth details and individual-level information, including residential data (complete residential postcode) to calculate the exact distances to the nearest PCS. Another strength of this research was the use of my novel genotype classification.

I did not take into account wind direction/speed, topography, or dispersion of bioaerosol. However, I increased the radius surrounding the RCS to 4 km to account for both wind direction and speed and small numbers for shorter distances in this research. Additionally, I was unable to take into consideration background bioaerosol concentrations (including seasonal change) or other potential anthropogenetic bioaerosol sources (such as intensive farming and sewage treatment operations). Thorough exposure assessment techniques should be taken into consideration for future studies. I geocoded PCSs using postcodes given by Environment Agencies, which might not accurately match the PCSs' exact locations, so there is a possibility of misclassification exposure. A further factor that would likely affect exposure is the site type or size of the PCS, such as whether it is an open windrow (located outside) where emissions are uncontained or an in-vessel (located indoors), where emissions are contained. However, there was limited information on these factors from the environmental agencies. Bioaerosols generated by PCS are expected to reach background levels within a few miles (Williams *et al.*, 2019). There was a small percentage of pwCF living within small distances around PCS; therefore, I was unable to examine the risk in a short-distance area or within 250 metres. As a result, I was unable to explore risks where concentrations of bioaerosol are predicted to be high (Deacon *et al.*, 2009; Douglas *et al.*, 2016a; Pankhurst *et al.*, 2011; Pearson *et al.*, 2015). The UK Environment Agency's existing risk assessment limit (250m) may be supported or adjusted by household measurements or by closely observing the few pwCF who live nearby PCS.

Like every registry study, there is a chance of bias, especially from unmeasured variables. The small increase in #IVdays, a small drop in ppFEV$_1$ and association with fungal and bacterial illness among pwCF living within 4 km of a PCS may be due to residual confounding since those who live nearby may differ from those who live farther away. I may have missed some

potentially important clinical parameters as I only looked at the most common bacterial and fungal species. For instance, future analyses should consider antifungal therapy. I am also aware that inhaled antibiotics are a proxy marker of chronic infection of pseudomonas and that diabetes mellitus is associated with worse outcomes, but due to the lack of data available for the analysis of this project, I was unable to use chronic infection of pseudomonas and adjust the effect of diabetes mellitus and inhaled antibiotic treatment in my regression models. The frequency and method of sampling (e.g., cough swab, sputum and bronchoscopy) will have differed, and the results of the sputum were based on UK CF Registry guidelines (for example, a larger percentage of children would have submitted cough swabs). It would be beneficial to conduct additional research using longitudinal sampling and molecular microbiological techniques to:

i) Improve understanding of the fungi that pwCF experience

ii) Find out whether pwCF who are exposed to more bioaerosols have higher fungal burden densities.

iii) Examine the pwCF respiratory mycobiome's fungal diversity and its relation to seasonal variables.

Future studies should consider wind speed and direction, topography, or bioaerosol dispersion, especially when exploring the effect of bioaerosol from a composting site on the disease severity in pwCF. Wind speed plays an important role in dispersing bioaerosols over longer distances, while wind direction determines any potential impact on nearby areas. Studies have shown the relevance of wind speed and direction in bioaerosol dispersion, underscoring the importance of considering these factors when estimating the potential impacts of bioaerosol on the health of people living near composting sites (Douglas *et al.*, 2017b; Douglas *et al.*, 2016b; Wang *et al.*, 2023; Wei *et al.*, 2020). Dispersion models help to calculate bioaerosol concentrations with high temporal and spatial resolution, resulting in an exposure assignment that is more precise and accurate (Douglas *et al.*, 2017b). Therefore, dispersion models are advised in future research for bioaerosol exposure assignments.

4.4.3 Conclusions and recommendations

This work contributes to the body of evidence of a relationship between living close to a PCS (a proxy for exposure to bioaerosol) and disease severity in pwCF in the UK. To further

evaluate the clinical significance of these results, studies using longitudinal sampling, molecular techniques for identifying and quantifying pathogens in sputum, and better bioaerosol exposure measures that account for bioaerosol dispersion are recommended. If my findings are confirmed, the association found in this study may have significant implications for the effects on pwCF living nearby PCS, their living surroundings, the therapeutic clinical advice given to them, their living conditions, and how PCS are permitted and regulated.

5 Chapter V: Overall discussion

My doctoral thesis aimed to better understand the impact of environmental factors on CF in the UK. To do this, I have looked at the effect of NO_2 and $PM_{2.5}$ on disease severity including lung function among children and adults with CF in Great Britain and at the effect of $PM_{2.5}$ exposure on longitudinal changes in lung function among a cohort of pwCF with a residential postcode in Greater London overa period of ten years (2008-2017). I also aimed to gain a better understanding of the influence of residential location relative to PCS on disease severity and the prevalence of fungal and bacterial infections in pwCF. Throughout this work, I used high-quality clinical data from the UK CF Registry, which is believed to capture over 99% of pwCF in the UK, combined with high-resolution surface models to estimate average concentrations of exposure to NO_2 and $PM_{2.5}$ at the residential postcode of pwCF. Several covariates were considered and adjusted for in the regression analyses to explore the effect of air pollutants and bioaerosols on disease severity in CF. These variables were age, sex, height, weight, BMI, genotype, *P. aeruginosa* infection sputum culture positive, deprivation (Townsend index) and country of residency.

5.1.1.1 Directed Acyclic Graphs (DAGs)

Directed Acyclic Graphs (DAGs) are recognized as valuable tools in epidemiology. They offer benefits such as visualizing causal relationships, identifying confounders and mediators, and enhancing conceptual clarity (Digitale *et al.*, 2022; Greenland *et al.*, 1999). While their application can be highly informative, their effectiveness hinges on a clear understanding of the underlying causal structure with minimal unmeasured confounding. In this study, I explored the potential association between $PM_{2.5}$, NO_2, and bioaerosols and CF severity, measured by $ppFEV_1$ and PEx. The complexity of the research design, with three exposures and two outcomes, presented a challenge for constructing a clear and interpretable DAG. A DAG with multiple exposure and outcome variables can become visually cluttered, obscuring the direction of the arrows and hindering the comprehension of relationships between the factors. Consequently, obtaining clear and interpretable results from such a DAG plot would have been difficult. Hence, I did not explicitly consider the use of a DAG plot in my thesis. However, it is important to acknowledge the strengths that DAGs offer in outlining causal pathways and potential sources of bias (VanderWeele & Robins, 2007). In this study, the focus

was on exploring the potential link between air pollution and CF severity, rather than definitively establishing causality. Given this exploratory nature and the research objectives, robust traditional statistical methods with adjustments for known confounders were employed alongside sensitivity analyses. This approach provided a solid methodological foundation for investigating the initial associations under study.

Although the results of my three core thesis chapters highlight the complexity of studies to quantify associations between environmental factors and disease severity, I believe that my work offers convincing evidence of the detrimental effects of some environmental exposures on the health of pwCF's lungs, in particular $PM_{2.5}$, with a more pronounced effect in adults with CF. The observed significant association between higher exposure to $PM_{2.5}$ levels and the decline in $ppFEV_1$ trajectories (***chapter IV***) underscores the importance of minimising environmental exposure to $PM_{2.5}$ concentrations for protecting lung function in pwCF. In addition, the association between $PM_{2.5}$ exposure and #IVdays suggests that higher $PM_{2.5}$ concentrations might contribute to a higher number of annual days of IV antibiotics. I also found that CF adults who lived close to a PCS had more days on IV antibiotics, which highlights the significant role that bioaerosol exposure may play in influencing respiratory health, and the need for more regular monitoring and/or earlier intervention with additional treatment if there is a decline. My findings show the importance of monitoring and addressing ambient air pollution levels in areas where pwCF are living. The findings of this investigation consistently showed a stronger association between exposure to $PM_{2.5}$ and disease severity in adults with CF compared to children with CF across the three key analyses of my thesis. The more pronounced effects of $PM_{2.5}$ exposure in CF adults imply that they may be more susceptible to its detrimental effects, either because of prolonged cumulative exposure, or a higher number of exacerbations or underlying comorbidities. This result highlights the need for earlier intervention to reduce the impact of $PM_{2.5}$ exposure, particularly in CF adults.

This study highlights the detrimental impact of $PM_{2.5}$ on pwCF, underscoring the need for targeted therapies and methods of $PM_{2.5}$ exposure reduction. It is important to translate these findings into practice to improve care and health outcomes in pwCF. By understanding the specific risks associated with $PM_{2.5}$ exposure, clinicians may develop personalised treatment

plans and preventive measures to improve the health outcomes of pwCF. The findings of this study emphasise the need for efficient research translation initiatives.

The study offers evidence of the adverse impacts of $PM_{2.5}$ in this CF population and provides essential data to support the CF Trust's work to promote better air quality, increase awareness, and implement preventive measures. The CF Trust might make use of these findings to develop personalised therapies, educational materials, and awareness campaigns to reduce $PM_{2.5}$ exposure among pwCF. The study may also help the CF Trust in its advocacy work for stricter air quality standards and regulations to protect the respiratory health of pwCF.

In this project, I devised a novel genotype classification in collaboration with clinicians and researchers, which is a significant contribution to the field. This new classification is in line with the most recent *CFTR* modulators, which may help clinicians improve the overall health of pwCF. In all the analyses I did for this project, I used this new genotype classification. As a sensitivity analysis, I also used the classification of existing genotypes to see if there were any variations in the results. The findings of both analyses were not different, demonstrating the robustness of the new classification.

Studying environmental-health associations is challenging due to reliance on proxy data. Though I had access to comprehensive clinical data on pwCF and high-resolution modelled exposure data available, capturing accurate and individual-specific exposure data is difficult due to pwCF mobility in the environment and indoor exposure variability. Proxy data may provide valuable information, but they might not fully capture the variability and specificity of an individual's exposures (Park & Kwan, 2017). To address this limitation, future studies could explore including personal monitoring methods. A number of studies have employed personal monitors to measure participants' exposure to air pollution (Heydon & Chakraborty, 2020; Kortoçi *et al.*, 2022; Xie *et al.*, 2021).

Sensitivity analyses and subgroup analyses are essential for ensuring the reliability and robustness of the results and for making them more generalizable (Deshmukh, 2021). I employed a number of sensitivity and subgroup analyses in this thesis. My results for the subgroup trajectory analysis (***section 3.3.5***) were similar to the primary findings for adults and

pwCF with a severe genotype, giving more thorough evidence of the relationship between $PM_{2.5}$ and lung function decline in pwCF.

Some outcomes are better proxies of CF health than others, but none of them provide a full picture given the complexity of particulates, fungi and bacteria involved in CF. I believe that the change in the rate of $ppFEV_1$ is a better marker of disease progression in general and in the context of exposure to air pollutants (Caley *et al.*, 2021). The rate of change in $ppFEV_1$ provides essential information about the progression of the CF disease over time but requires data from multiple years. Monitoring the decline in $ppFEV_1$ over time could help identify high-risk pwCF earlier who may experience significant disease progression. In research by Marsteller et al., using changes in FEV_1 data to create a CF risk disease progression score tool, it was reported that pwCF were ten times more likely to have severe disease progression if their score was at least two on assessment (Marsteller *et al.*, 2019) (Marsteller *et al.*, 2019). This finding highlights the importance of using changes in FEV_1 over time. Monitoring the rate of change in FEV_1 is essential in the context of exposure to air pollution. Exposure to air pollutants causes airway inflammation and oxidative stress, which accelerates the loss of lung function in pwCF (Blayac *et al.*, 2022). Air pollution has been shown to have harmful impacts on respiratory health (Bălă *et al.*, 2021; Manisalidis *et al.*, 2020; Strassmann *et al.*, 2021). Therefore, monitoring the rate of change in $ppFEV_1$ in pwCF exposed to air pollutants is important for disease progression and identifying pwCF who may require early interventions to reduce the harmful effects of air pollutant exposures.

The findings of my study did not show an association between NO_2 exposure and disease severity in either children or adults with CF in the UK. This lack of association may be due to the ecological fallacy, which is common in environmental health studies (Idrovo, 2011; Loney & Nagelkerke, 2014). It is essential to understand this fallacy and its implications for interpreting and contextualising my research findings. In the present study, while no significant association was found between NO_2 exposure and disease severity in the pwCF, it does not necessarily imply that there is no association at the individual level. The ecological fallacy may occur in environmental health studies for a variety of reasons, such as proxy measure of exposure or some unmeasured effect modifiers. This absence of a significant association might have been due to some unmeasured cofounder, like smoking history or

other comorbidities that were not included in my investigation. A number of studies have reported a rapid decline in lung function among smokers (Kopp *et al.*, 2016; Oates *et al.*, 2020; Ong *et al.*, 2017). Future research should include smoking data to better understand the independent effects of NO_2 exposure on disease severity in pwCF. Another explanation could be the timing and duration of the NO_2 exposure, as the cross-sectional study only used a single-point exposure and may not have captured the crucial exposure windows. Compared to single-point exposure, long-term exposure to NO_2 in a longitudinal study may have a different effect. Even though there was no statistically significant relationship at the population level, future studies should include all potential effect modifiers at the individual level and personal monitoring exposure techniques.

The emergence of home monitoring equipment, combined with health and fitness trackers for personalised monitoring, holds significant promise for pwCF (Rebecca Jane *et al.*, 2020; Warrier & Sood, 2021). Home monitoring technology enables individuals to check and keep track of their health in their own homes in real time (Liu *et al.*, 2016). Health and fitness trackers have become more popular recently, along with home monitoring technology (Henriksen *et al.*, 2018). These wrist-worn gadgets or smartphone add-ons provide a variety of features including activity tracking, heart rate monitoring, sleep analysis, daily activities, diet and overall health. Nevertheless, despite these advancements, we still know very little about how environmental factors impact CF (Rebecca Jane *et al.*, 2020). More research is needed to bridge this knowledge gap and explore the complex relationships between CF disease severity and environmental factors using the data collected from these advanced monitoring techniques.

My study's findings highlight the need for more research and provide a compelling case for international collaborations to further explore these associations on a larger scale. Such collaborations can provide numerous advantages, including larger sample sizes and enhanced genotype diversity, which can result in more robust and generalizable findings. By collaborating with researchers from different nations, it would be feasible to combine data from several populations, leading to more statistically significant relationships and enabling subgroup analysis based on various genetic mutations or demographic characteristics with higher power. Additionally, these collaborations promote the exchange of knowledge,

expertise, and resources, which may help standardise methodologies for data collection across all sites.

5.1.2 Strength of the evidence on the relationship between air pollution exposure, bioaerosol and CF outcomes

The strength of evidence on the relationship between air pollution exposure, bioaerosol and CF outcomes, particularly $PM_{2.5}$ exposure, was robust and compelling based on the findings of my doctoral thesis. The study utilized high-quality clinical data from the UK CF Registry, capturing over 99% of pwCF in the UK, combined with high-resolution exposure models for NO_2 and $PM_{2.5}$ and including all active RCSs in 2016 across the UK. The observed significant associations between higher $PM_{2.5}$ exposure levels and adverse outcomes, including the decline in $ppFEV_1$ and an increased #IVdays, provided a solid foundation for understanding the impact of air pollutants on CF severity.

The evidence was further strengthened by the comprehensive consideration of potential confounding factors, including age, sex, genotype, height, weight, *P. aeruginosa* sputum culture positive and deprivation in the regression analyses. The use of sensitivity and subgroup analyses increased the reliability and robustness of the results, ensuring that the associations withstood rigorous scrutiny.

While the study did not find a significant association with NO_2 exposure, it is essential to acknowledge the potential limitations, such as the ecological fallacy. Unlike $PM_{2.5}$, which can be relatively constant across geographic areas (de Jesus *et al.*, 2019), NO_2 exposure can vary considerably at the individual level due to factors like traffic density and proximity to specific sources. This spatial variation might weaken the association observed in our analysis using area-level NO_2 data. To strengthen the evidence on NO_2, future studies could incorporate individual-level exposure data, potentially through residential monitoring or wearable air quality sensors. This approach would allow for a more robust assessment of the relationship between NO_2 and CF severity.

In summary, the robust methodology, large-scale registry data, and rigorous statistical analyses contributed to a high level of confidence in the observed associations between $PM_{2.5}$ exposure and adverse CF disease severity. However, ongoing research and international

collaborations can further strengthen the evidence base, addressing the remaining research gaps and refining our understanding of the intricate relationship between air pollution and CF disease severity.

5.1.3 Recommendations for the CF community

The findings of my research are aligned with one of CF's top ten research priorities, which was developed with the James Lind Alliance in January 2017 (James Lind alliance PSP., 2022). My finding of a significant association between $PM_{2.5}$ exposure and lung function and bioaerosol and #IVdays for PEx is aligned with the priority research question "What are the effective ways of simplifying the treatment burden of people with Cystic Fibrosis?". Reducing $PM_{2.5}$ exposure and living away from composting sites may significantly reduce the burden of treatment in pwCF. This was confirmed by Bo et *al*. who found that lower $PM_{2.5}$ levels were associated with better lung function (Bo *et al.*, 2021). They found that every 5 $\mu g/m^3$ decrease in $PM_{2.5}$ per year is associated with 19.9 mL/year (95%CI: 17.4; 22.4) in FEV_1 in 133,119 adults (aged 18 or older) in Taiwan. By matching our research efforts with the goals stated by pwCF, the researcher can address the most important problems and make a significant and meaningful impact on pwCF's lives and those of their families.

Postcodes are used to estimate exposure at the area level in many health studies, particularly in environmental epidemiology. Therefore, an accurate and complete postcode is important for an accurate and true exposure assignment. Missing postcodes not only affect true exposure assignment but also affect the accuracy and representativeness of the data, leading to inaccurate results. It also reduces the power of the study and limits the generalizability of its findings. Researchers studying the impact of air pollutants on disease severity in pwCF rely on reliable postcode data to assess exposure accurately and draw informed conclusions. Although the quality of the postcode data in the UK CF Registry was good overall, I identified some gaps and inconsistencies. Therefore, the UK CF registry should ensure the accuracy and completeness of the postcode field for all the pwCF in the database. This will not only improve the data's completeness and accuracy but will also increase its usefulness and value to a researcher who will use UK CF Registry data to produce robust studies. Additionally, the UK CF Registry should collect more detailed and reliable data on smoking in the Registry.

5.1.4 Key public health implications of my findings

1. Protecting vulnerable populations such as CF from air pollution: My findings highlighted the detrimental impact of $PM_{2.5}$ exposure on lung function and exacerbations in adults with CF. This reinforced the urgent need for stricter air quality regulations and policies, particularly in areas where there are high-risk individuals such as CF populations and empowers them with strategies to minimize their exposure. A recent survey showed that 31% of adults with CF had moved to somewhere with better air quality or a better environment for pwCF (Jamie E. *et al.*, 2023). Public health efforts should prioritize reducing ambient $PM_{2.5}$ levels and informing the CF community about the risks and strategies to minimize exposure.

2. Location matters for CF health: I identified an association between living near composting sites and increased IV antibiotic use in pwCF. This underscores the potential impact of bioaerosol exposure on respiratory health. It emphasizes the importance of considering bioaerosol risks when setting up particularly open-air windrow composting facilities and public awareness campaigns for those living near such sites.

3. Empowering the CF community: Public health efforts should actively engage the CF community in awareness campaigns and educational resources. This will empower individuals to make informed decisions about their environment and manage their health proactively.

4. Continued research for better interventions: My research contributed to the small body of evidence linking environmental factors with disease severity in CF. However, ongoing research is crucial to fully understand these complex relationships. Future studies should focus on longitudinal investigations, personal exposure monitoring, and exploring emerging biomarkers for earlier detection and intervention.

In conclusion, my research highlighted the public health imperative of prioritizing environmental considerations in CF care, advocating for cleaner air, and implementing preventive measures to safeguard the respiratory health of individuals with CF.

5.1.5 Implications for clinicians and pwCF

My research offered valuable insights for both clinicians and pwCF in navigating the potential impact of environmental factors on their health. Here is a breakdown of the key implications:

5.1.5.1 For Clinicians

- **Enhanced risk assessment:** Consider environmental exposures, particularly $PM_{2.5}$ and proximity to composting sites, as potential contributors to disease progression and exacerbations, especially in adults with CF.
- **Personalized treatment plans:** Integrate environmental factors into treatment plans. Consider interventions like air purifiers and avoidance strategies for high-risk CF individuals, such as limiting outdoor activities during peak pollution hours if possible.
- **Patient education and advocacy:** Increase awareness among patients about the risks of air pollution and bioaerosols and provide comprehensive and up-to-date resources and strategies for minimizing exposure. Advocate for improved air quality standards and responsible siting of composting facilities.
- **Collaboration with environmental health professionals:** Partner with environmental health experts to gain deeper insights into local air quality and potential bioaerosol risks, tailoring advice to specific geographic contexts.
- **Health lifestyle** pwCF should continue to be encouraged to have a healthy lifestyle that includes eating a balanced diet, getting regular exercise, and taking their prescription medications as prescribed, as recommended by the European Cystic Fibrosis Society (ECFS) (Southern *et al.*, 2023). These actions can help decrease the impact of air pollution while also enhancing overall respiratory health.

In summary, clinicians should integrate environmental considerations into their patient care strategies, tailoring interventions based on individual exposure profiles. Empowering pwCF to actively engage in managing their environmental exposures will contribute to improved health outcomes and an enhanced quality of life.

5.1.5.2 For pwCF

- **Understanding your environment:** First and foremost, pwCF should be aware of the air pollution levels outdoors and indoors, around their residences, schools and places of employment and the location of potential bioaerosol sources like composting sites. Use air quality apps and consider personal monitoring devices to understand your individual exposure levels.
- pwCF who live close to a composting site should consider taking extra safety measures to reduce their exposure to bioaerosols, such as wearing face masks and staying indoors during the peak composting times, when feasible.
- **Minimize exposure:** Taking proactive steps could help reduce exposure to harmful pollutants. This might include using air purifiers at home, wearing masks during times of high air pollution, and limiting activities outdoors during peak pollution hours. Use an indoor air purifier that may efficiently reduce airborne pollutants and other potentially dangerous contaminants in the indoor air (Dubey *et al.*, 2021). Studies have demonstrated that these purifiers may significantly reduce indoor $PM_{2.5}$ air pollution (Chen *et al.*, 2022; Vijayan *et al.*, 2015). Pay particular attention to their symptoms when air pollution levels are high, avoid areas with high pollution at specific times of the day; and carefully plan the route when going outside. Wearing a suitable face mask may help reduce exposure to air pollution when these other prevention measures can be used. Using real-time mobile apps may help pwCF avoid high-pollution areas and routes. These applications give real-time information and alerts regarding air pollution, helping pwCF make an informed decision about going outdoors. pwCF may take preventative measures by limiting their exposure to harmful air pollutants, which will lessen the effect of air pollution on their respiratory health.
- **Communicate with your doctor:** Discuss your environmental concerns with your doctor and work together to develop a personalized plan to manage environmental risks. Inform them of any changes in your environment or lifestyle that might impact your health.
- **Advocate for yourself:** If possible, actively participate in patient advocacy groups and raise awareness about the impact of environmental factors on CF. Support efforts to improve air quality regulations and responsible siting of bioaerosol sources

5.1.6 Directions for the future work

I used 2016 clinical data to study the influence of NO_2 and $PM_{2.5}$ on disease severity in pwCF in GB, as this was the most up-to-date available at the time of the analysis of this research. There was also a fair amount of data processing involved; therefore, I conducted the analysis only for one year for this project. Running this analysis over multiple years would have enabled me to further validate my results and possibly refine the assessment of health effects. Given the substantial differences in air pollution levels and pwCF's access to care before and after SARS-CoV-2, it will be interesting to evaluate the same study using clinical data for pre-2020 and post-2020 (Bakola *et al.*, 2022; Davies, 2020; Munir *et al.*, 2021).

Future research should concentrate on examining the impact of NO_2 on lung function trajectories in pwCF, which would help in the development of preventative measures for improving the health of pwCF. The effects of NO_2 on disease severity in people with other respiratory disorders, such as asthma, COPD, or respiratory infections, are well-documented (Fu *et al.*, 2022; Huangfu & Atkinson, 2020; Zheng *et al.*, 2021). Future studies on NO_2 will help us better understand the complex relationship between the effects of individual and multiple pollutants on the severity of disease in pwCF.

The lung function trajectory analysis was conducted on a relatively small sample only in pwCF with a residential postcode in Greater London due to the availability of exposure data. This model assigns $PM_{2.5}$ values specifically for the Greater London area. Future research should use a similar high-resolution model for air pollution exposure covering the whole of the UK. With the use of such a model, one would be able to expand upon the research and include pwCF from all around the country. This expansion would enhance the generalizability and representativeness of the findings, allowing for a more comprehensive understanding of the relationship between air pollution and lung function trajectory in pwCF at a national level.

In my trajectory analysis, I used 2008 - 2017 clinical data. Adding the latest years of clinical data to study the $pppFEV_1$ trajectories over a longer period may have allowed the identification and consistency of any distinct pattern in the trajectories of $ppFEV_1$ in pwCF (Agusti & Faner, 2019).

Future research should concentrate on examining how other air pollutants, such as O_3 and PM_{10}, impact the disease severity in pwCF, which would help in the development of preventative measures for improving the health of pwCF. The effects of O_3 and PM_{10} on disease severity in people with other respiratory disorders, such as asthma, COPD, or respiratory infections, are well-documented (Bălă *et al.*, 2021; Cortes *et al.*, 2023; Fu *et al.*, 2022; Medina-Ramón *et al.*, 2006). Future studies on O_3 and PM_{10} pollutants will help us better understand the complex relationship between the effects of individual and multiple pollutants on the severity of disease in pwCF.

My analyses on the impact of air pollutants and bioaerosol exposure on disease severity in pwCF were conducted at three different geographies, mainly due to the availability of exposure data. The first analysis, exploring the impact of air pollutants on disease severity using 2016 data, was conducted at the GB level. The $ppFEV_1$ trajectory analysis focused on Greater London, while my composting analysis included the entire UK. To ensure comprehensive and comparable results, it is recommended to obtain high-resolution exposure data at the UK level. This would allow all three analyses to be conducted at the same scale, providing a more comprehensive understanding of the relationship between air pollutants, bioaerosol exposure and disease severity in pwCF.

Future studies should consider the mobility patterns of pwCF as well, as estimating air pollutant exposure based just on residential location will not adequately capture the actual exposure to air pollutants. It would miss the dynamic air pollution exposure based on a person's daily mobility behaviours. Understanding pwCF mobility patterns, including time spent at locations other than their residence, will improve the accuracy and relevance of study results. Numerous studies have highlighted the significance of taking individual mobility patterns, behaviour, and spatial variations in exposure to air pollutants into account in environmental epidemiological studies (Marquet *et al.*, 2023; Nyhan *et al.*, 2016; Yoo *et al.*, 2021). Future research on the effects of air pollution on respiratory health should take into account how people move around. This will help to understand exposure profiles more thoroughly.

In examining the influence of bioaerosols from composting sites on disease severity, I would perform also a subset of analysis in composting sites which are open windrows. The bioaerosol

emissions from this kind of facility are more harmful than those from in-vessel systems (enclosed) as the emissions from open windrow facilities are not confined hence the effect on the disease severity in pwCF living near PCS would be the largest (Robertson *et al.*, 2019).
The Lung Clearance Index (LCI), one of the emerging physiologic biomarkers, may be used in future studies to evaluate the health of pwCF's airways because it is a sensitive indicator of early changes, especially in mild lung disease, a repeatable and useful measure of CF airway disease, and more effective than conventional lung function tests at detecting early airway disease in children with CF (Horsley, 2009). There is a limitation to LCI as it is only superior to FEV_1 in mild lung disease, so the use of LCI across the full age and disease severity spectrum of lung disease remains uncertain (Ellemunter & Steinkamp, 2019).

Smoking has been shown to have detrimental effects on CF lung disease, PEx and a rapid decline in lung function (Oates *et al.*, 2020). Though not many pwCF smoke (Mc Ewan *et al.*, 2012) studies show worse outcomes in CF children and second-hand smoke exposure (Benjamin *et al.*, 2019; Kopp *et al.*, 2016). Therefore, it is important to measure and take smoking into account in future CF population studies. Furthermore, research should aim to develop targeted interventions to promote smoking cessation among pwCF. Oates *et al.* found that cessation of smoking reduces the odds of PEx by 17% ($p< 0.001$), a 0.7% increase in $ppFEV_1$ ($p< 0.001$), and a 1% increase in the BMI percentile ($p < 0.001$) in the first year of cessation among children and adolescents with CF (Oates *et al.*, 2021). In summary, future studies should emphasise the importance of adjusting for smoking in assessing disease severity in CF and investigate the potential benefits of smoking cessation in this CF population. By addressing these, we can improve our understanding of the impact of smoking on CF and develop effective strategies to promote healthier outcomes for the CF population.

The microbiological data I used predominantly was taken from the registry which represents a single respiratory culture at the time of annual review, whereas an alternative (and perhaps more accurate/relevant) indicator, may have been to use a definition for ‘chronic infection’, ie. based on a percentage number of positive cultures over the prior year (Keating *et al.*, 2021)r. Future research could consider data on inhaled antibiotics for chronic *P. aeruginosa* infections, as these are surrogate markers of chronic *P. seudomonas* infection in pwCF.

DM and pancreatic insufficiency are regarded as indicators of disease severity in CF therefore, they should be taken into account in future studies (Coderre *et al.*, 2021; Singh & Schwarzenberg, 2017). Weather conditions are important because of their significant impact on both pwCF and air quality. They affect clinical conditions of CF, as reported by the effects of temperature (Collaco *et al.*, 2011; Psoter *et al.*, 2016) and season (Psoter *et al.*, 2016) on *P. aeruginosa* infection in the lungs. Therefore, these meteorological variables should be considered to account for their effect in future studies.

During the COVID-19 pandemic, there has been a considerable reduction in outdoor air pollution levels, mostly as a result of decreased industrial activity and lower traffic emissions as a result of lockdown measures (Jephcote *et al.*, 2021). Reduced availability of routine healthcare services, including non-emergency medical treatment, was another effect of this pandemic on healthcare access. These have had a huge impact on changing people's habits to stay indoors more often than go outdoors (Ferreira & Barros, 2022). Therefore, indoor air pollution should also be a target of future research to understand the impact of indoor air quality on disease severity in pwCF.

5.1.7 Conclusions

In conclusion, my doctoral thesis contributes to an understanding of the relationship between air pollutants, bioaerosols and disease severity in pwCF. The key findings include the detrimental effects of $PM_{2.5}$ exposure on lung function decline, disease severity, and the prevalence of bacterial and fungal infections in individuals with CF. Notably, the study underscores the importance of considering $PM_{2.5}$ exposure reduction strategies for protecting lung function in CF patients, especially in adults. The observed associations, supported by high-quality clinical data from the UK CF Registry, reveal the complex relationship between environmental exposures and CF health outcomes. My findings highlight the need for preventative measures to reduce air pollution exposure and protect the respiratory health of vulnerable populations, such as those with CF. By focusing on the environmental factors that contribute to disease severity, we may improve the quality of life and health outcomes for pwCF and people with other respiratory diseases. Future studies should continue to investigate the effects of air pollutants while taking into consideration the key data covered in the section of recommendations above. By taking into consideration all these factors, one

may gain a more comprehensive understanding of the influence of air pollutants on health outcomes and devise efficient public health measures to lower exposure to air pollution in pwCF.

6 References

Aatamila, M. *et al.* (2011) Odour annoyance and physical symptoms among residents living near waste treatment centres. *Environmental Research.* 111 (1), 164-170.

Adams, J. *et al.* (2005) How accurate are Townsend Deprivation Scores as predictors of self-reported health? A comparison with individual level data. *Journal of Public Health.* 27 1, 101-6.

Agusti, A. & Faner, R. (2019) Lung function trajectories in health and disease. *The Lancet. Respiratory medicine.* 7 (4), 358-364.

Allen MP. (1997) The problem of multicollinearity. In: Allen, M.P. (ed.) *Understanding Regression Analysis*. Boston, MA, Springer US, pp. 176-180.

Amin, R. *et al.* (2010) The effect of chronic infection with Aspergillus fumigatus on lung function and hospitalization in patients with cystic fibrosis. *Chest.* 137 (1), 171-6.

Ashworth, M. *et al.* (2021) Spatio-temporal associations of air pollutant concentrations, GP respiratory consultations and respiratory inhaler prescriptions: a 5-year study of primary care in the borough of Lambeth, South London. *Environmental Health.* 20 (1), 54.

Avol, E. L. *et al.* (2001) Respiratory effects of relocating to areas of differing air pollution levels. *American Journal of Respiratory and Critical Care Medicine.* 164 (11), 2067-72.

Ayyagari, V. N. *et al.* (2007) Effects of Nitrogen Dioxide on the Expression of Intercellular Adhesion Molecule-1, Neutrophil Adhesion, and Cytotoxicity: Studies in Human Bronchial Epithelial Cells. *Inhalation Toxicology.* 19 (2), 181-194.

Bakare, N. *et al.* (2003) Prevalence of Aspergillus fumigatus and other fungal species in the sputum of adult patients with cystic fibrosis. *Mycoses.* 46 (1-2), 19-23.

Bakola, M. *et al.* (2022) The impact of COVID-19 lockdown on air pollution in Europe and North America: a systematic review. *European Journal of Public Health.* 32 (6), 962-968.

Bălă, G. P. *et al.* (2021) Air pollution exposure-the (in)visible risk factor for respiratory diseases. *Environmental Science and Pollution Research Journal.* 28 (16), 19615-19628.

Baldwin, N. *et al.* (2015) Factors affecting pollutant concentrations in the near-road environment. *Atmospheric Environment.* 115, 223-235.

Balfour-Lynn, I. M. & King, J. A. (2022) CFTR modulator therapies - Effect on life expectancy in people with cystic fibrosis. *Paediatric Respiratory Reviews.* 42, 3-8.

Balmain, A. (2023) Air pollution's role in the promotion of lung cancer. *Nature.* 616 (7955), 35-36.

Barone-Adesi, F. *et al.* (2015) Long-Term Exposure to Primary Traffic Pollutants and Lung Function in Children: Cross-Sectional Study and Meta-Analysis. *PloS One.* 10 (11), e0142565.

Barr, H. L. *et al.* (2011) Association between socioeconomic status, sex, and age at death from cystic fibrosis in England and Wales (1959 to 2008): cross sectional study. *British Medical Journal.* 343, d4662.

Başaran, A. E. *et al.* (2021) Association Between Cystic Fibrosis Severity Markers and CFTR Genotypes in Turkish Children. *Turkish thoracic journal.* 22 (6), 426-431.

Bates, D. *et al.* (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software.* 67 (1), 1 - 48.

Beelen, R. *et al.* (2008) Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environmental Health Perspectives.* 116 (2), 196-202.

Beevers, S. D. *et al.* (2012) One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions. *Atmospheric environment.* 59, 47-58.

Beevers, S. D. *et al.* (2013) Air pollution dispersion models for human exposure predictions in London. *Journal of Exposure Science & Environmental Epidemiology.* 23 (6), 647-653.

Benjamin, T. K. *et al.* (2019) Secondhand smoke alters arachidonic acid metabolism and inflammation in infants and children with cystic fibrosis. *Thorax.* 74 (3), 237.

Berger, M. (2002) *Inflammatory mediators in cystic fibrosis lung disease*, *Allergy and Asthma Proceedings*, OceanSide Publications.

Bhagirath, A. Y. *et al.* (2016) Cystic fibrosis lung environment and Pseudomonas aeruginosa infection. *BMC Pulmonary Medicine.* 16 (1), 174.

Blayac, M. *et al.* (2022) The Impact of Air Pollution on the Course of Cystic Fibrosis: A Review. *Frontiers in Physiology.* 13, 908230.

Blomberg, A. *et al.* (1999) Persistent airway inflammation but accommodated antioxidant and lung function responses after repeated daily exposure to nitrogen dioxide. *American Journal of Respiratory and Critical Care Medicine.* 159 (2), 536-43.

Bo, Y. *et al.* (2021) Reduced ambient PM2.5, better lung function, and decreased risk of chronic obstructive pulmonary disease. *Environment International.* 156, 106706.

Bobadilla, J. L. *et al.* (2002) Cystic fibrosis: a worldwide analysis of CFTR mutations—correlation with incidence data and application to screening. *Human Mutation.* 19 (6), 575-606.

Bolund, A. C. S. *et al.* (2017) The effect of organic dust exposure on long-term change in lung function: a systematic review and meta-analysis. *Occupational and Environmental Medicine.* 74 (7), 531.

Boniardi, L. *et al.* (2019) Is a Land Use Regression Model Capable of Predicting the Cleanest Route to School? *Environments.* 6 (8), 90.

Bos, B. *et al.* (2023) Prenatal exposure to air pollution is associated with structural changes in the neonatal brain. *Environment International.* 174, 107921.

Bowatte, G. *et al.* (2017) Traffic-related air pollution exposure over a 5-year period is associated with increased risk of asthma and poor lung function in middle age. *The European respiratory journal.* 50 (4).

Bowe, B. *et al.* (2019) Burden of cause-specific mortality associated with PM2. 5 air pollution in the United States. *The Journal of the American Medical Association network open.* 2 (11), e1915834-e1915834.

Brauer, M. *et al.* (2006) Traffic-related air pollution and otitis media. *Environmental Health Perspectives.* 114 (9), 1414-8.

Brauer, M. *et al.* (2007) Air pollution and development of asthma, allergy and infections in a birth cohort. *The European respiratory journal.* 29 (5), 879-88.

Breuer, O. *et al.* (2018) Predicting disease progression in cystic fibrosis. *Expert Review of Respiratory Medicine.* 12 (11), 905-917.

Breza-Boruta, B. (2016) The assessment of airborne bacterial and fungal contamination emitted by a municipal landfill site in Northern Poland. *Atmospheric Pollution Research.* 7 (6), 1043-1052.

British society for ecological medicine. (2018) *The Health Effects of Waste Incinerators.* *https://www.bsem.org.uk/pages/156-the-health-effects-of-waste-incinerators* [Accessed 28th June 20203].

Brown, J. S. *et al.* (2001) Regional deposition of coarse particles and ventilation distribution in healthy subjects and patients with cystic fibrosis. *Journal of aerosol medicine : the official journal of the International Society for Aerosols in Medicine.* 14 4, 443-54.

Brugha, R. *et al.* (2018) Outdoor air pollution and cystic fibrosis. *Paediatric Respiratory Reviews.* 28, 80-86.

Bunger, J. *et al.* (2000) Health complaints and immunological markers of exposure to bioaerosols among biowaste collectors and compost workers. *Occup Environ Med.* 57 (7), 458-64.

Bunger, J. *et al.* (2007) A 5-year follow-up study on respiratory disorders and lung function in workers exposed to organic dust from composting plants. *International Archives of Occupational and Environmental Health.* 80 (4), 306-12.

Burgel, P.-R. *et al.* (2016) Aspergillus fumigatus in the cystic fibrosis lung: pros and cons of azole therapy. *Infection and drug resistance.* 9, 229-238.

Burnett, R. *et al.* (2018) Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences.* 115 (38), 9592-9597.

Burnham, K. P. & Anderson, D. R. (2003) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* Springer New York.

Caley, L. *et al.* (2021) Average rate of lung function decline in adults with cystic fibrosis in the United Kingdom: Data from the UK CF registry. *Journal of Cystic Fibrosis.* 20 (1), 86-90.

Caruana, D. J. (2011) Detection and analysis of airborne particles of biological origin: present and future. *Analyst.* 136 (22), 4641-52.

CF Europe. (2023) *Living with CF* https://www.cf-europe.eu/ [Accessed 26th March 2023].

Chamberlain, R. C. *et al.* (2023) Health effects of low emission and congestion charging zones: a systematic review. *Lancet Public Health.* 8 (7), e559-e574.

Chang, C. H. *et al.* (2022) Ambient Air Pollution and Risk for Stroke Hospitalization: Impact on Susceptible Groups. *Toxics.* 10 (7).

Chen, C. F. *et al.* (2022) Efficacy of HEPA Air Cleaner on Improving Indoor Particulate Matter 2.5 Concentration. *International Journal of Environmental Research and Public Health.* 19 (18).

Chen, J. & Hoek, G. (2020) Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis. *Environment International.* 143, 105974.

Chen, Q. *et al.* (2021) A review of cystic fibrosis: Basic and clinical aspects. *Animal Models and Experimental Medicine.* 4 (3), 220-232.

Chen, T. M. *et al.* (2007) Outdoor air pollution: nitrogen dioxide, sulfur dioxide, and carbon monoxide health effects. *The American journal of the medical sciences.* 333 (4), 249-56.

Chiappini, E. *et al.* (2014) Bacterial lung infections in cystic fibrosis patients: an update. *Pediatric Infectious Disease Journal.* 33 (6), 653-4.

Clark, N. A. *et al.* (2010) Effect of early life exposure to air pollution on development of childhood asthma. *Environmental Health Perspectives.* 118 (2), 284-90.

Clauss, M. (2015) Particle size distribution of airborne micro-organisms in the environment-A review. *Landbauforschung Volkenrode.* 65, 77-100.

Clinical and Functional Translation of CFTR (CFTR2). (2022a) *Baltimore (MD):The Clinical and Functional Translation of CFTR (CFTR2), US Cystic Fibrosis Foundation, Johns Hopkins University, the Hospital for Sick Children*. https://www.cff.org/sites/default/files/2021-12/Study-VX18-445-106-Eligible-Mutations-June-2019.pdf [Accessed 29th September 2022].

Clinical and Functional Translation of CFTR (CFTR2). (2022b) https://cftr2.org/ [Accessed 03rd March 2022].

Coderre, L. *et al.* (2021) The Potential Causes of Cystic Fibrosis-Related Diabetes. *Frontiers in Endocrinology.* 12, 702823.

Cohen, A. J. *et al.* (2017) Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet.* 389 (10082), 1907-1918.

Coleman, N. C. *et al.* (2020) Fine particulate matter exposure and cancer incidence: analysis of SEER cancer registry data from 1992–2016. *Environmental Health Perspectives.* 128 (10), 107004.

Collaco, J. M. *et al.* (2010) Quantification of the relative contribution of environmental and genetic factors to variation in cystic fibrosis lung function. *The Journal of pediatrics.* 157 (5), 802-7.e1-3.

Collaco, J. M. *et al.* (2011) Effect of temperature on cystic fibrosis lung disease and infections: a replicated cohort study. *PloS One.* 6 (11), e27784.

Collaco, J. M. *et al.* (2008) Interactions Between Secondhand Smoke and Genes That Affect Cystic Fibrosis Lung Disease. *The Journal of the American Medical Association.* 299 (4), 417-24.

Cooper, N. *et al.* (2019) Inequalities in exposure to the air pollutants PM2.5 and NO2 in Australia. *Environmental Research Letters.* 14 (11).

Cortes, T. R. *et al.* (2023) Short-term association between ambient air pollution and cardio-respiratory mortality in Rio de Janeiro, Brazil. *PloS One.* 18 (2), e0281499.

Coverstone, A. M. & Ferkol, T. W. (2020) Early Diagnosis and Intervention in Cystic Fibrosis: Imagining the Unimaginable. *Frontiers in Pediatrics.* 8, 608821.

Crawford, J. A. *et al.* (2015) Indicators of airborne fungal concentrations in urban homes: understanding the conditions that affect indoor fungal exposures. *Science of the Total Environment.* 517, 113-24.

Cromwell, E. A. *et al.* (2023) Cystic fibrosis prevalence in the United States and participation in the Cystic Fibrosis Foundation Patient Registry in 2020. *Journal of Cystic Fibrosis.* 22 (3), 436-442.

Cuthbertson, L. *et al.* (2021) The fungal airway microbiome in cystic fibrosis and non-cystic fibrosis bronchiectasis. *Journal of cystic fibrosis* 20 (2), 295-302.

Cystic Fibrosis Foundation (n.d.) *About Cystic Fibrosis.* https://www.cff.org/What-is-CF/About-Cystic-Fibrosis/ [Accessed 26th March 2023].

Cystic Fibrosis Trust (2024) Strategic Research Centre: CF-EpiNet Harnessing data to improve lives. https://www.imperial.ac.uk/nhli/research/centres-and-initiatives/strategic-research-centre-cystic-fibrosis-epinet/. [Accessed 23 May. 2019].

da Silveira Fleck, A. *et al.* (2021) Environmental and Occupational Short-Term Exposure to Airborne Particles and FEV(1) and FVC in Healthy Adults: A Systematic Review and Meta-Analysis. *International Journal of Environmental Research and Public Health.* 18 (20).

Dauchet, L. *et al.* (2018) Short-term exposure to air pollution: associations with lung function and inflammatory markers in non-smoking, healthy adults. *Environment international.* 121, 610-619.

Davies, J. (2020) The coronavirus pandemic has forced rapid changes in care protocols for cystic fibrosis. *Nature.* 583 (7818), S15.

De Boeck, K. & Amaral, M. D. (2016) Progress in therapies for cystic fibrosis. *The Lancet Respiratory Medicine.* 4 (8), 662-674.

De Boeck, K. *et al.* (2014) Efficacy and safety of ivacaftor in patients with cystic fibrosis and a non-G551D gating mutation. *Journal of Cystic Fibrosis.* 13 (6), 674-680.

de Boer, K. *et al.* (2011) Exacerbation frequency and clinical outcomes in adult patients with cystic fibrosis. *Thorax.* 66 (8), 680-5.

de Hoogh, K. *et al.* (2016) Development of West-European PM2.5 and NO2 land use regression models incorporating satellite-derived and chemical transport modelling data. *Environmental Research.* 151, 1-10.

de Jesus, A. L. *et al.* (2019) Ultrafine particles and PM2.5 in the air of cities around the world: Are they representative of each other? *Environment International.* 129, 118-135.

Deacon, L. *et al.* (2009) Endotoxin emissions from commercial composting activities. *Environmental Health.* 8 (1), S9.

DEFRA. (2014) *Department of Environment Food and Rural Affairs, 2014. Waste Exemption: T23 Aerobic Composting and Associated Prior Treatment.* https://www.gov.uk/guidance/waste-exemption-t23-aerobic-composting-and-associated-prior-treatment [Accessed 03rd March. 2022].

DEFRA. (2021) *Air Quality Targets in the Environment Act* https://uk-air.defra.gov.uk/library/air-quality-targets [Accessed 19th June 2023].

DEFRA. (2022) *UK Statistics on Waste UK statistics on waste - GOV.UK* (www.gov.uk) [Accessed 15th June 2023].

DEFRA. (2023a) *National statistics Emissions of air pollutants in the UK – Nitrogen oxides (NOx).* https://www.gov.uk/government/statistics/emissions-of-air-pollutants/emissions-of-air-pollutants-in-the-uk-nitrogen-oxides-nox#major-emission-sources-for-nitrogen-oxides-in-the-uk [Accessed 18th June 2023].

DEFRA. (2023b) *National statistics Emissions of air pollutants in the UK – Particulate matter (PM10 and PM2.5).* https://www.gov.uk/government/statistics/emissions-of-air-pollutants/emissions-of-air-pollutants-in-the-uk-particulate-matter-pm10-and-pm25 [Accessed 19th June 2023].

DEFRA. (2023c) *National statistics Nitrogen dioxide (NO2).* https://www.gov.uk/government/statistics/air-quality-statistics/ntrogen-dioxide [Accessed 18th June 2023].

Delfino, E. *et al.* (2019) Respiratory Fungal Diseases in Adult Patients With Cystic Fibrosis. *Clinical Medicine Insights: Circulatory, Respiratory and Pulmonary Medicine.* 13, 1179548419849939.

Delfino, R. J. *et al.* (2008) Personal and ambient air pollution exposures and lung function decrements in children with asthma. *Environmental Health Perspectives.* 116 (4), 550-8.

Desbiolles, A. *et al.* (2018) Cancer incidence in adults living in the vicinity of nuclear power plants in France, based on data from the French Network of Cancer Registries. *International Journal of Cancer.* 142 (5), 899-909.

Deshmukh, M. (2021) Sensitivity and Subgroup Analyses. In: Patole, S. (ed.) *Principles and Practice of Systematic Reviews and Meta-Analysis*. Cham, Springer International Publishing, pp. 89-97.

Di, Q. *et al.* (2017) Air Pollution and Mortality in the Medicare Population. *The New England journal of medicine.* 376 (26), 2513-2522.

Digitale, J. C. *et al.* (2022) Tutorial on directed acyclic graphs. *Journal of Clinical Epidemiology.* 142, 264-267.

Doiron, D. *et al.* (2019) Air pollution, lung function and COPD: results from the population-based UK Biobank study. *The European respiratory journal.* 54 (1).

Dominici, F. *et al.* (2006) Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *The Journal of the American Medical Association.* 295 (10), 1127-1134.

Don, B. S. & Christopher Hooper, G. (2013) Pulmonary exacerbations as indicators of progression of lung disease in young children with CF. *Thorax.* 68 (7), 608.

Dos Anjos, F. *et al.* (2022) Pancreatic Insufficiency in a Child with p.Gly542* and c.2657+5G>A Heterozygote CFTR: A Case Report. *Clinical Medicine & Research.* 20 (1), 46-51.

Douglas, P. *et al.* (2016a) Respiratory hospital admission risk near large composting facilities. *International journal of hygiene and environmental health.* 219 (4-5), 372-9.

Douglas, P. *et al.* (2017a) Estimating Particulate Exposure from Modern Municipal Waste Incinerators in Great Britain. *Environmental Science & Technology.* 51 (13), 7511-7519.

Douglas, P. *et al.* (2017b) Use of dispersion modelling for Environmental Impact Assessment of biological air pollution from composting: Progress, problems and prospects. *Waste Management.* 70, 22-29.

Douglas, P. *et al.* (2016b) Sensitivity of predicted bioaerosol exposure from open windrow composting facilities to ADMS dispersion model parameters. *Journal of Environmental Management.* 184, 448-455.

Douwes, J. *et al.* (2003) Bioaerosol health effects and exposure assessment: progress and prospects. *Annals of Occupational Hygiene.* 47 (3), 187-200.

Drumm, M. L. *et al.* (2012) Genetic variation and clinical heterogeneity in cystic fibrosis. *Annual Review of Pathology.* 7, 267-282.

Dubey, S. *et al.* (2021) Assessing effectiveness of air purifiers (HEPA) for controlling indoor particulate pollution. *Heliyon.* 7 (9), e07976.

Durda-Masny, M. *et al.* (2021) The determinants of survival among adults with cystic fibrosis-a cohort study. *Journal of Physiological Anthropology.* 40 (1), 19.

Earnest, A. *et al.* (2020) Lung function over the life course of paediatric and adult patients with cystic fibrosis from a large multi-centre registry. *Scientific Reports.* 10 (1), 17421.

EEA (2021) *Healthy environment, healthy lives: how the environment influences health and well-being in Europe.* https://www.eea.europa.eu/themes/air/health-impacts-of-air-pollution, [Accessed 10th May 2022].

Eeftens, M. *et al.* (2016) Development of land use regression models for nitrogen dioxide, ultrafine particles, lung deposited surface area, and four other markers of particulate matter pollution in the Swiss SAPALDIA regions. *Environmental Health.* 15 (1), 53.

Eitan, K. *et al.* (2014) Factors associated with FEV₁ decline in cystic fibrosis: analysis of the ECFS Patient Registry. *European Respiratory Journal.* 43 (1), 125.

Elbarbary, M. *et al.* (2020) Ambient air pollution, lung function and COPD: cross-sectional analysis from the WHO Study of AGEing and adult health wave 1. *BMJ open respiratory research.* 7 (1).

Elborn, J. S. (2016) Cystic fibrosis. *Lancet.* 388 (10059), 2519-2531.

Eleonora Fichera. *et al.* (2023) *Low Emission Zones improve air quality, physical health and mental well-being. Zenodo - DOI: 10.5281/zenodo.8360285 [Accessed 20th Feb 2024].*

Ellemunter, H. & Steinkamp, G. (2019) Lung clearance index to detect the efficacy of Aztreonam lysine inhalation in patients with cystic fibrosis and near normal spirometry - A single-centre feasibility study. *PloS One.* 14 (9), e0221673.

Elpis, H. *et al.* (2023) Risk factors for FEV1 decline in European patients with CF: data from the European Cystic Fibrosis Society Patient Registry (ECFSPR). *ERJ open research.* 00449-2022.

Environment Agency. (2016) *Risk assessments for your environmental permit. https://www.gov.uk/guidance/risk-assessments-for-your-environmental-permit [Accessed 03rd March. 2022].*

Environment Agency. (2018) *Bioaerosol Monitoring at Regulated Facilities - Use of M9: RPS 209.* https://www.gov.uk/government/publications/bioaerosol-monitoring-at-regulated-facilities-use-of-m9-rps-209 [Accessed 20th Feb 2022].

Er, B. *et al.* (2020) Factors associated with severe lung disease in an adult population with cystic fibrosis: a single-center experience. *Turkish journal of medical sciences.* 50 (4), 945-952.

ESRI. (2018) ArcGIS Desktop: Release 10.7.1 Redlands, CA: Environmental Systems Research Institute (ESRI).

European Commission. (2012) *European Commission, 2012. Environment. Landfill of waste.* http://ec.europa.eu/environment/waste/landfill_index.htm [Accessed 21st September 2020].

Evangelopoulos, D. *et al.* (2021) Personal exposure to air pollution and respiratory health of COPD patients in London. *The European respiratory journal.* 58 (1).

Farhat, S. C. L. *et al.* (2013) Ozone is associated with an increased risk of respiratory exacerbations in patients with cystic fibrosis. *Chest.* 144 (4), 1186-1192.

Ferrari, U. *et al.* (2012) Influence of air pressure, humidity, solar radiation, temperature, and wind speed on ambulatory visits due to chronic obstructive pulmonary disease in Bavaria, Germany. *International Journal of Biometeorology.* 56 (1), 137-143.

Ferreira, A. & Barros, N. (2022) COVID-19 and Lockdown: The Potential Impact of Residential Indoor Air Quality on the Health of Teleworkers. *International Journal of Environmental Research and Public Health.* 19 (10).

Fiore, M. *et al.* (2020) Correctors modify the bicarbonate permeability of F508del-CFTR. *Scientific Reports.* 10 (1), 8440.

Fischer, A. J. *et al.* (2021) Sustained Coinfections with Staphylococcus aureus and Pseudomonas aeruginosa in Cystic Fibrosis. *American Journal of Respiratory and Critical Care Medicine.* 203 (3), 328-338.

Flannigan, B. *et al.* (1991) Allergenic and toxigenic micro-organisms in houses. *Society for Applied Bacteriology Symposium Series.* 20, 61s-73s.

Flight, W. G. *et al.* (2014) The effect of the weather on pulmonary exacerbations and viral infections among adults with cystic fibrosis. *International Journal of Biometeorology.* 58 (9), 1845-1851.

Fu, Y. *et al.* (2022) Association and interaction of O3 and NO2 with emergency room visits for respiratory diseases in Beijing, China: a time-series study. *BMC Public Health.* 22 (1), 2265.

Gan, W. Q. *et al.* (2012) Association of long-term exposure to community noise and traffic-related air pollution with coronary heart disease mortality. *American Journal of Epidemiology.* 175 (9), 898-906.

Gao, N. *et al.* (2019) Short-term effects of ambient air pollution on chronic obstructive pulmonary disease admissions in Beijing, China (2013–2017). *International Journal of Chronic Obstructive Pulmonary Disease.* 14, 297.

Garcia-Clemente, M. *et al.* (2020) Impact of Pseudomonas aeruginosa Infection on Patients with Chronic Inflammatory Airway Diseases. *Journal of Clinical Medicine Research.* 9 (12).

Gauderman, W. J. *et al.* (2004) The Effect of Air Pollution on Lung Development from 10 to 18 Years of Age. *New England Journal of Medicine.* 351 (11), 1057-1067.

Gehring, U. *et al.* (2015) Particulate Matter Composition and Respiratory Health: The PIAMA Birth Cohort Study. *Epidemiology.* 26 (3).

Gilligan, P. H. (2014) Infections in patients with cystic fibrosis: diagnostic microbiology update. *Clinics in Laboratory Medicine.* 34 (2), 197-217.

Goeminne, P. C. *et al.* (2013) Impact of air pollution on cystic fibrosis pulmonary exacerbations: a case-crossover analysis. *Chest.* 143 (4), 946-954.

Goss, C. H. & Burns, J. L. (2007) Exacerbations in cystic fibrosis·1: Epidemiology and pathogenesis. *Thorax.* 62 (4), 360-367.

Goss, C. H. *et al.* (2004) Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine.* 169 (7), 816-21.

Gourd, E. (2022) New evidence that air pollution contributes substantially to lung cancer. *The Lancet Oncology.* 23 (10), e448.

Govan, J. R. *et al.* (2007) Evolving epidemiology of Pseudomonas aeruginosa and the Burkholderia cepacia complex in cystic fibrosis lung infection. *Future Microbiology.* 2 (2), 153-64.

Greater London Authority. (2024) Inner London Ultra Low Emission Zone Expansion One Year Report. https://www.london.gov.uk/programmes-strategies/environment-and-climate-change/environment-and-climate-change-publications/inner-london-ultra-low-emission-zone-expansion-one-year-report [Accessed 4th April 2024].

Greenland, S. *et al.* (1999) Causal diagrams for epidemiologic research. *Epidemiology.* 10 (1), 37-48.

Grehn, C. *et al.* (2020) Urban Life as Risk Factor for Aspergillosis. *Frontiers in Cellular and Infection Microbiology.* 10.

Gulliver, J. *et al.* (2013) Development and back-extrapolation of NO2 land use regression models for historic exposure assessment in Great Britain. *Environmental Science & Technology.* 47 (14), 7804-11.

Gurram, S. *et al.* (2019) Agent-based modeling to estimate exposures to urban air pollution from transportation: Exposure disparities and impacts of high-resolution data. *Computers, Environment and Urban Systems.* 75, 22-34.

Hajat, A. *et al.* (2015) Socioeconomic Disparities and Air Pollution Exposure: a Global Review. *Current environmental health reports.* 2 (4), 440-50.

Hajna, S. *et al.* (2023) Associations between residential greenspace exposure and mortality in 4 645 581 adults living in London, UK: a longitudinal study. *The Lancet Planetary Health.* 7 (6), e459-e468.

Hambach, R. *et al.* (2012) Work-related health symptoms among compost facility workers: a cross-sectional study. *Archives of Public Health.* 70 (1), 13.

Hanigan, I. C. *et al.* (2019) All-cause mortality and long-term exposure to low level air pollution in the '45 and up study' cohort, Sydney, Australia, 2006–2015. *Environment International.* 126, 762-770.

Hannah Ritchie. & Max Roser. (2022) *Outdoor Air Pollution.* https://ourworldindata.org/outdoor-air-pollution. [Accessed 21st May 2020].

Haq, I. *et al.* (2022) Precision Medicine Based on CFTR Genotype for People with Cystic Fibrosis. *Pharmacogenomics and Personalized Medicine.* 15, 91-104.

Harron, K. *et al.* (2017) Challenges in administrative data linkage for research. *Big Data & Society.* 4 (2), 2053951717745678.

Harun, S. N. *et al.* (2016) A systematic review of studies examining the rate of lung function decline in patients with cystic fibrosis. *Paediatric Respiratory Reviews.* 20, 55-66.

Hatziagorou, E. *et al.* (2023) Risk factors for FEV1 decline in European patients with CF: data from the European Cystic Fibrosis Society Patient Registry (ECFSPR). *ERJ open research.* 00449-2022.

Hatziagorou, E. *et al.* (2020) Changing epidemiology of the respiratory bacteriology of patients with cystic fibrosis–data from the European cystic fibrosis society patient registry. *Journal of Cystic Fibrosis.* 19 (3), 376-383.

Hatzis, C. *et al.* (2006) Ambient particulate matter exhibits direct inhibitory effects on oxidative stress enzymes. *Environmental Science & Technology.* 40 (8), 2805-11.

Hayes, R. B. *et al.* (2020) PM2.5 air pollution and cause-specific cardiovascular disease mortality. *International Journal of Epidemiology.* 49 (1), 25-35.

He, D. *et al.* (2017) Association between particulate matter 2.5 and diabetes mellitus: A meta-analysis of cohort studies. *Journal of Diabetes Investigation.* 8 (5), 687-696.

Heijerman, H. G. M. *et al.* (2019) Efficacy and safety of the elexacaftor plus tezacaftor plus ivacaftor combination regimen in people with cystic fibrosis homozygous for the F508del mutation: a double-blind, randomised, phase 3 trial. *The Lancet.* 394 (10212), 1940-1948.

Henriksen, A. *et al.* (2018) Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of Medical Internet Research.* 20 (3), e110.

Herr, C. E. W. *et al.* (2003) Effects of bioaerosol polluted outdoor air on airways of residents: a cross sectional study. *Occupational and Environmental Medicine.* 60 (5), 336-342.

Heydon, J. & Chakraborty, R. (2020) Can portable air quality monitors protect children from air pollution on the school run? An exploratory study. *Environmental Monitoring and Assessment.* 192 (3), 195.

Hoek, G. *et al.* (2008) A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment.* 42 (33), 7561-7578.

Hoek, G. *et al.* (2013) Long-term air pollution exposure and cardio- respiratory mortality: a review. *Journal of Environmental Health.* 12 (1), 43.

Hoffmann, C. *et al.* (2022) Asthma and COPD exacerbation in relation to outdoor air pollution in the metropolitan area of Berlin, Germany. *Respiratory Research.* 23 (1), 64.

Hoffmeyer, F. *et al.* (2014) Prevalence of and relationship between rhinoconjunctivitis and lower airway diseases in compost workers with current or former exposure to organic dust. *Annals of Agricultural and Environmental Medicine.* 21 (4), 705-11.

Holgate, S. *et al.* (2016) *Every breath we take: The lifelong impact of air pollution, Report of a working party*, Royal College of Physicians.

Hoo, Z. H. *et al.* (2018) Do cystic fibrosis centres with the lowest FEV1 still use the least amount of intravenous antibiotics? A registry-based comparison of intravenous antibiotic use among adult CF centres in the UK. *Journal of Cystic Fibrosis.* 17 (3), 360-367.

Horsley, A. (2009) Lung clearance index in the assessment of airways disease. *Respiratory Medicine.* 103 (6), 793-9.

Hou, D. *et al.* (2020) Associations of long-term exposure to ambient fine particulate matter and nitrogen dioxide with lung function: A cross-sectional study in China. *Environment International.* 144, 105977.

Huangfu, P. & Atkinson, R. (2020) Long-term exposure to NO2 and O3 and all-cause and respiratory mortality: A systematic review and meta-analysis. *Environment International.* 144, 105998.

Ibrahim, A. J. & Abdul, R. (2017) The Cystic Fibrosis Airway Microbiome and Pathogens. In: Dinesh, S. (ed.) *Progress in Understanding Cystic Fibrosis*. Rijeka, IntechOpen, pp. Ch. 3.

Idrovo, A. J. (2011) Three criteria for ecological fallacy. *Environmental Health Perspectives.* 119 (8), A332.

Illek, B. *et al.* (1999) Defective function of the cystic fibrosis-causing missense mutation G551D is recovered by genistein. *American Journal of Physiology-Cell Physiology.* 277 (4), C833-C839.

Imperial College London (2024) A new MRC-funded doctoral training programme (DTP) in Multidisciplinary Science in Health at Imperial College. https://www.imperial.ac.uk/multisci-mrc-dtp/. [Accessed 23 May. 2019].

Jacobsen, G. *et al.* (2008) Longitudinal lung function decline and wood dust exposure in the furniture industry. *European Respiratory Journal.* 31 (2), 334.

James Lind alliance PSP. (2022) *Cystic Fibrosis Top 10.*https://www.jla.nihr.ac.uk/priority-setting-partnerships/cystic-fibrosis/top-10-priorities.htm [Accessed 21st June 2023].

Jamie E. *et al.* (2023) The Financial Costs of Cystic Fibrosis. https://www.cysticfibrosis.org.uk/sites/default/files/2023-01/CFT%20final%20report.pdf. [Accessed 4th April 2024].

Jassal, M. S. *et al.* (2013) Effect of residential proximity to major roadways on cystic fibrosis exacerbations. *International Journal of Environmental Health Research.* 23 (2), 119-31.

Jephcote, C. *et al.* (2021) Changes in air quality during COVID-19 'lockdown' in the United Kingdom. *Environmental Pollution.* 272, 116011.

Jia, H. *et al.* (2021) PM2.5-induced pulmonary inflammation via activating of the NLRP3/caspase-1 signaling pathway. *Environmental Toxicology.* 36 (3), 298-307.

Jiang, X. Q. *et al.* (2016) Air pollution and chronic airway diseases: what should people know and do? *The Journal of Thoracic Disease.* 8 (1), E31-40.

Junge, S. *et al.* (2016) Factors associated with worse lung function in cystic fibrosis patients with persistent Staphylococcus aureus. *PloS One.* 11 (11), e0166220.

Kamdar, O. *et al.* (2008) Air pollution induces enhanced mitochondrial oxidative stress in cystic fibrosis airway epithelium. *FEBS Letters.* 582 (25-26), 3601-6.

Kanchongkittiphon, W. *et al.* (2015) Indoor environmental exposures and exacerbation of asthma: an update to the 2000 review by the Institute of Medicine. *Environmental Health Perspectives.* 123 (1), 6-20.

Keating, C. L. *et al.* (2021) Pseudomonas aeruginosa Susceptibility Patterns and Associated Clinical Outcomes in People with Cystic Fibrosis following Approval of Aztreonam Lysine for Inhalation. *Antimicrobial Agents and Chemotherapy.* 65 (3).

Kelly, F. *et al.* (2011) The London low emission zone baseline study. *Research Report (Health Effects Institute).* (163), 3-79.

Kelly, F. J. (2003) Oxidative stress: its role in air pollution and adverse health effects. *Journal of Occupational and Environmental Medicine.* 60 (8), 612-6.

Kelly, F. J. & Fussell, J. C. (2011) Air pollution and airway disease. *Clinical and Experimental Allergy.* 41 (8), 1059-1071.

Keogh, R. H. *et al.* (2020) The changing demography of the cystic fibrosis population: forecasting future numbers of adults in the UK. *Scientific Reports.* 10 (1), 10660.

Kerem, B. *et al.* (1989) Identification of the cystic fibrosis gene: genetic analysis. *Science.* 245 (4922), 1073-80.

Kerem, E. *et al.* (2014) Factors associated with FEV₁ decline in cystic fibrosis: analysis of the ECFS Patient Registry. *European Respiratory Journal.* 43 (1), 125.

Kiedrowski, M. R. & Bomberger, J. M. (2018) Viral-Bacterial Co-infections in the Cystic Fibrosis Respiratory Tract. *Frontiers in Immunology.* 9, 3067.

Kim, J. *et al.* (2016) Living near nuclear power plants and thyroid cancer risk: A systematic review and meta-analysis. *Environment International.* 87, 42-8.

Kim, J. J. (2004) Ambient air pollution: health hazards to children. *Pediatrics.* 114 (6), 1699-1707.

Kim, S. Y. & Song, I. (2017) National-scale exposure prediction for long-term concentrations of particulate matter and nitrogen dioxide in South Korea. *Environmental Pollution.* 226, 21-29.

Konstan, M. W. *et al.* (2007) Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. *The Journal of Pediatrics.* 151 (2), 134-9, 139.e1.

Kopp, B. T. *et al.* (2016) The Impact of Secondhand Smoke Exposure on Children with Cystic Fibrosis: A Review. *International Journal of Environmental Research and Public Health.* 13 (10).

Kortoçi, P. *et al.* (2022) Air pollution exposure monitoring using portable low-cost air quality sensors. *Smart Health.* 23, 100241.

Kraemer, R. *et al.* (2006) Progression of pulmonary hyperinflation and trapped gas associated with genetic and environmental factors in children with cystic fibrosis. *Respiratory Research.* 7 (1), 138.

Kuhn, D. M. & Ghannoum, M. A. (2003) Indoor mold, toxigenic fungi, and Stachybotrys chartarum: infectious disease perspective. *Clinical Microbiology Reviews.* 16 (1), 144-172.

Künzi, L. *et al.* (2015) Toxicity of aged gasoline exhaust particles to normal and diseased airway epithelia. *Scientific Reports.* 5.

Kyte, L. & Wells, C. (2010) Variations in life expectancy between rural and urban areas of England, 2001–07. *Health Statistics Quarterly.* 46 (1), 27-52.

Kyung, S. Y. & Jeong, S. H. (2020) Particulate-Matter Related Respiratory Diseases. *Tuberculosis and Respiratory Diseases (Seoul).* 83 (2), 116-121.

Lacey, J. & Dutkiewicz, J. (1994) Bioaerosols and occupational lung disease. *Journal of Aerosol Science.* 25 (8), 1371-1404.

Lago-Peñas, S. *et al.* (2021) The impact of socioeconomic position on non-communicable diseases: what do we know about it? *Perspectives in public health.* 141 (3), 158-176.

Lalley, P. M. (2013) The aging respiratory system—Pulmonary structure, function and neural control. *Respiratory Physiology & Neurobiology.* 187 (3), 199-210.

Lee, G. H. *et al.* (2020a) Effects of Indoor Air Purifiers on Children with Asthma. *Yonsei Medical Journal.* 61 (4), 310-316.

Lee, K. K. *et al.* (2020b) Adverse health effects associated with household air pollution: a systematic review, meta-analysis, and burden estimation study. *The Lancet Global Health.* 8 (11), e1427-e1434.

Lelieveld, J. *et al.* (2015) The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature.* 525 (7569), 367-71.

Leung, G. J. *et al.* (2020) Variation in lung function and nutritional decline in cystic fibrosis by genotype: An analysis of the Canadian cystic fibrosis registry. *Journal of Cystic Fibrosis.* 19 (2), 255-261.

Li, N. *et al.* (2022) Association of PM2.5 and PM10 with Acute Exacerbation of Chronic Obstructive Pulmonary Disease at lag0 to lag7: A Systematic Review and Meta-Analysis. *COPD: Journal of Chronic Obstructive Pulmonary Disease.* 19 (1), 243-254.

Li, S. *et al.* (2020) Long-term exposure to PM2.5 and Children's lung function: a dose-based association analysis. *Journal of Thoracic Disease.* 12 (10), 6379-6395.

Li, T. *et al.* (2018) Fine particulate matter (PM(2.5)): The culprit for chronic lung diseases in China. *Chronic Diseases and Translational Medicine.* 4 (3), 176-186.

Liang, H. *et al.* (2023) Association of outdoor air pollution, lifestyle, genetic factors with the risk of lung cancer: A prospective cohort study. *Environmental Research.* 218, 114996.

Ling, S. H. & van Eeden, S. F. (2009) Particulate matter air pollution exposure: role in the development and exacerbation of chronic obstructive pulmonary disease. *International Journal of Chronic Obstructive Pulmonary Disease.* 4, 233.

Liou, T. G. *et al.* (2001) Predictive 5-Year Survivorship Model of Cystic Fibrosis. *American Journal of Epidemiology.* 153 (4), 345-352.

Liu, L. *et al.* (2016) Smart homes and home health monitoring technologies for older adults: A systematic review. *International Journal of Medical Informatics.* 91, 44-59.

Liu, Q. *et al.* (2017) Effect of exposure to ambient PM(2.5) pollution on the risk of respiratory tract diseases: a meta-analysis of cohort studies. *Journal of Biomedical Materials Research.* 31 (2), 130-142.

London Online. (2023) *London Boroughs.*https://www.londononline.co.uk/boroughs/ [Accessed 30th June 2023].

Loney, T. & Nagelkerke, N. J. (2014) The individualistic fallacy, ecological studies and instrumental variables: a causal interpretation. *Emerging Themes in Epidemiology.* 11 (1), 18.

Lopes-Pacheco, M. (2020) CFTR Modulators: The Changing Face of Cystic Fibrosis in the Era of Precision Medicine. *Frontiers in Pharmacology.* 10.

Lu, X. *et al.* (2021) Airway hyperresponsiveness development and the toxicity of PM2. 5. *Environmental Science and Pollution Research.* 28 (6), 6374-6391.

Lukacs, G. L. & Verkman, A. S. (2012) CFTR: folding, misfolding and correcting the ΔF508 conformational defect. *Trends in Molecular Medicine.* 18 (2), 81-91.

Lyczak, J. B. *et al.* (2002) Lung infections associated with cystic fibrosis. *Clinical Microbiology Reviews.* 15 (2), 194-222.

Ma, X. *et al.* (2020) PyLUR: Efficient software for land use regression modeling the spatial distribution of air pollutants using GDAL/OGR library in Python. *Frontiers of Environmental Science & Engineering.* 14 (3), 44.

Macdougall, A. *et al.* (2023) Trajectories of early growth and subsequent lung function in cystic fibrosis: An observational study using UK and Canadian registry data. *Journal of Cystic Fibrosis.* 22 (3), 388-394.

Mack, S. M. *et al.* (2019) Respiratory Health Effects of Exposure to Ambient Particulate Matter and Bioaerosols. *Comprehensive Physiology.* 10 (1), 1-20.

Magee, L. C. *et al.* (2021) Managing Fungal Infections in Cystic Fibrosis Patients: Challenges in Clinical Practice. *Infection and Drug Resistance.* 14, 1141-1153.

Mahadeva, R. *et al.* (1998) Clinical outcome in relation to care in centres specialising in cystic fibrosis: cross sectional studyCommentary: management in paediatric and adult cystic fibrosis centres improves clinical outcome. *British Medical Journal.* 316 (7147), 1771-1775.

Manisalidis, I. *et al.* (2020) Environmental and Health Impacts of Air Pollution: A Review. *Frontiers in Public Health.* 8, 14.

Marquet, O. *et al.* (2023) Comparison of static and dynamic exposures to air pollution, noise, and greenness among seniors living in compact-city environments. *International Journal of Health Geographics.* 22 (1), 3.

Marsteller, N. L. *et al.* (2019) Cystic fibrosis patients at risk for disease progression marked by decline in FEV1% predicted: development of the cystic fibrosis risk of disease progression score. *The Journal of Thoracic Disease.* 11 (12), 5557-5565.

Maturu, V. & Agarwal, R. (2015) Prevalence of A spergillus sensitization and allergic bronchopulmonary aspergillosis in cystic fibrosis: systematic review and meta-analysis. *Clinical and Experimental Allergy.* 45 (12), 1765-1778.

Mc Ewan, F. A. *et al.* (2012) The prevalence of "risky behaviour" in adults with cystic fibrosis. *Journal of Cystic Fibrosis.* 11 (1), 56-8.

McGarry, M. E. *et al.* (2019) The demographics of adverse outcomes in cystic fibrosis. *Journal of Pediatric Pulmonology.* 54 Suppl 3 (Suppl 3), S74-s83.

Medina-Ramón, M. *et al.* (2006) The effect of ozone and PM10 on hospital admissions for pneumonia and chronic obstructive pulmonary disease: a national multicity study. *American Journal of Epidemiology.* 163 (6), 579-88.

Mei Zahav, M. *et al.* (2023) Disease severity of people with cystic fibrosis carrying residual function mutations: Data from the ECFS Patient Registry. *Journal of Cystic Fibrosis.* 22 (2), 234-247.

Mendell, M. J. *et al.* (2011) Respiratory and allergic health effects of dampness, mold, and dampness-related agents: a review of the epidemiologic evidence. *Environmental Health Perspectives.* 119 (6), 748-56.

Middleton, P. G. *et al.* (2019) Elexacaftor–Tezacaftor–Ivacaftor for Cystic Fibrosis with a Single Phe508del Allele. *New England Journal of Medicine.* 381 (19), 1809-1819.

Milanzi, E. B. *et al.* (2018) Air pollution exposure and lung function until age 16 years: the PIAMA birth cohort study. *European Respiratory Journal.* 52 (3), 1800218.

Milla, C. E. (2007) Nutrition and lung disease in cystic fibrosis. *Clinics in Chest Medicine.* 28 (2), 319-30.

Monoson, A. *et al.* (2023) Air pollution and respiratory infections: the past, present, and future. *Toxicological Sciences.* 192 (1), 3-14.

Morgenstern, V. *et al.* (2008) Atopic diseases, allergic sensitization, and exposure to traffic-related air pollution in children. *American Journal of Respiratory and Critical Care Medicine.* 177 (12), 1331-7.

Moss, R. B. *et al.* (2015) Efficacy and safety of ivacaftor in patients with cystic fibrosis who have an Arg117His-CFTR mutation: a double-blind, randomised controlled trial. *The Lancet. Respiratory medicine.* 3 (7), 524-33.

Müller, T. *et al.* (2006) Acute blood neutrophilia induced by short-term compost dust exposure in previously unexposed healthy individuals. *International Archives of Occupational and Environmental Health.* 79 (6), 477.

Munir, S. *et al.* (2021) Changes in Air Quality Associated with Mobility Trends and Meteorological Conditions during COVID-19 Lockdown in Northern England, UK. *Atmosphere.* 12 (4), 504.

Nagy, R. *et al.* (2022) Association of Body Mass Index With Clinical Outcomes in Patients With Cystic Fibrosis: A Systematic Review and Meta-analysis. *JAMA network open.* 5 (3), e220740.

Nasari, M. M. *et al.* (2016) A class of non-linear exposure-response models suitable for health impact assessment applicable to large cohort studies of ambient air pollution. *Air Quality, Atmosphere and Health.* 9 (8), 961-972.

Nelson, L. T. (2020) Cystic Fibrosis Lung Disease: An Overview. *Respiratory Care.* 65 (2), 233.

Nicholas, B. *et al.* (2017) Susceptibility to influenza virus infection of bronchial biopsies in asthma. *Journal of Allergy and Clinical Immunology.* 140 (1), 309-312.e4.

Nieuwenhuijsen, M. J. *et al.* (2018) Air Pollution, Noise, Blue Space, and Green Space and Premature Mortality in Barcelona: A Mega Cohort. *International Journal of Environmental Research and Public Health.* 15 (11).

Norbäck, D. *et al.* (2011) Lung function decline in relation to mould and dampness in the home: the longitudinal European Community Respiratory Health Survey ECRHS II. *Thorax.* 66 (5), 396-401.

Nyhan, M. *et al.* (2016) "Exposure Track"—The Impact of Mobile-Device-Based Mobility Patterns on Quantifying Population Exposure to Air Pollution. *Environmental Science & Technology.* 50 (17), 9671-9681.

Oates, G. R. *et al.* (2021) Cessation of smoke exposure improves pediatric CF outcomes: Longitudinal analysis of CF Foundation Patient Registry data. *Journal of Cystic Fibrosis.* 20 (4), 618-624.

Oates, G. R. *et al.* (2020) Tobacco smoke exposure and socioeconomic factors are independent predictors of pulmonary decline in pediatric cystic fibrosis. *Journal of Cystic Fibrosis.* 19 (5), 783-790.

Office for Health Improvement & Disparities. (2022) Air pollution: applying All Our Health [Accessed 17th June 2023].

Olivia Murrin. (2022) *Environmental influences on the pulmonary manifestations of Cystic Fibrosis.* PhD. Imperial College Lodnon, http://hdl.handle.net/10044/1/109119.

Ong, T. *et al.* (2017) Socioeconomic Status, Smoke Exposure, and Health Outcomes in Young Children With Cystic Fibrosis. *Pediatrics.* 139 (2).

ONS. (2013) *Office for National Statistics; Postcode Headcounts and Household Estimates - 2011 Census.* https://www.nomisweb.co.uk/census/2011/postcode_headcounts_and_household_estimates. [Accessed 22 Feb. 2022].

ONS. (2019) *ONS Postcode Directory (February 2019).* https://geoportal.statistics.gov.uk/items/ons-postcode-directory-february-2019 [Accessed 23 May. 2019].

ONS. (2020) *Office for National Statistics.. 2011 Census: Aggregate Data. [data collection]. UK Data Service. SN: 7427,* http://doi.org/10.5257/census/aggregate-2011-2 [Accessed 03 March 2022].

Ordnance Survey. (2022) *Ordnance Survey.* https://www.ordnancesurvey.co.uk/business-and-government/products/code-point-open.html [Accessed 08 March. 2022].

Orellano, P. *et al.* (2017) Effect of outdoor air pollution on asthma exacerbations in children and adults: Systematic review and multilevel meta-analysis. *PloS One.* 12 (3), e0174050.

Orellano, P. *et al.* (2020) Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. *Environment International.* 142, 105876.

Ortega-García, J. A. *et al.* (2012) Smoking prevention and cessation programme in cystic fibrosis: integrating an environmental health approach. *Journal of Cystic Fibrosis.* 11 (1), 34-9.

Pankhurst, L. J. *et al.* (2011) Spatial variations in airborne microorganism and endotoxin concentrations at green waste composting facilities. *International Journal of Hygiene and Environmental Health.* 214 (5), 376-383.

Pardo, M. *et al.* (2020) Particulate matter toxicity is Nrf2 and mitochondria dependent: the roles of metals and polycyclic aromatic hydrocarbons. *Chemical Research in Toxicology.* 33 (5), 1110-1120.

Park, Y. M. & Kwan, M.-P. (2017) Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place.* 43, 85-94.

Patel, A. R. *et al.* (2019) Diagnosing Allergic Bronchopulmonary Aspergillosis: A Review. *Cureus.* 11 (4), e4550-e4550.

Pearson, C. *et al.* (2015) Exposures and health outcomes in relation to bioaerosol emissions from composting facilities: a systematic review of occupational and community studies. *Journal of Toxicology and Environmental Health.* 18 (1), 43-69.

Persinger, R. L. *et al.* (2002) Molecular mechanisms of nitrogen dioxide induced epithelial injury in the lung. *Journal of Molecular and Cellular Biochemistry.* 234-235 (1-2), 71-80.

Pfeffer, P. E. *et al.* (2021) Air Pollution and Asthma: Mechanisms of Harm and Considerations for Clinical Interventions. *Chest.* 159 (4), 1346-1355.

Poli, P. *et al.* (2020) Transitory pancreatic insufficiency in cystic fibrosis children. *Journal of Cystic Fibrosis.* 19 (1), e5-e6.

Poore, T. S. *et al.* (2021) Fungal Infection and Inflammation in Cystic Fibrosis. *Pathogens.* 10 (5).

Population matters. (2018) *Waste in the UK and the world (2018).* https://populationmatters.org/sites/default/files/waste_in_the_uk_and_the_world.pdf. [Accessed 29th Sep 2020].

Psoter, K. J. *et al.* (2016) Association of meteorological and geographical factors and risk of initial Pseudomonas aeruginosa acquisition in young children with cystic fibrosis. *Epidemiology and Infection.* 144 (5), 1075-83.

Qibin, L. *et al.* (2020) The impact of PM2.5 on lung function in adults with asthma. *International Journal of Tuberculosis and Lung Disease.* 24 (6), 570-576.

Quanjer, P. H. *et al.* (2012) Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *The European respiratory journal.* 40 (6), 1324-43.

Qvist, T. *et al.* (2019) Seasonal fluctuation of lung function in cystic fibrosis: A national register-based study in two northern European populations. *Journal of Cystic Fibrosis.* 18 (3), 390-395.

R Core Team. (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rafaj, P. *et al.* (2018) Outlook for clean air in the context of sustainable development goals. *Global Environmental Change.* 53, 1-11.

Rajagopalan, S. & Landrigan, P. J. (2021) Pollution and the Heart. *New England Journal of Medicine.* 385 (20), 1881-1892.

Raju, S. V. *et al.* (2013) Cigarette smoke induces systemic defects in cystic fibrosis transmembrane conductance regulator function. *American Journal of Respiratory and Critical Care Medicine.* 188 (11), 1321-30.

Ramsay, K. A. *et al.* (2017) The changing prevalence of pulmonary infection in adults with cystic fibrosis: a longitudinal analysis. *Journal of Cystic Fibrosis.* 16 (1), 70-77.

Ramsay, K. A. *et al.* (2016) Infection in cystic fibrosis: impact of the environment and climate. *Expert Review of Respiratory Medicine.* 10 (5), 505-19.

Ramsey, B. W. *et al.* (2011) A CFTR Potentiator in Patients with Cystic Fibrosis and the G551D Mutation. *New England Journal of Medicine.* 365 (18), 1663-1672.

Ranganathan, S. C. *et al.* (2013) Geographical differences in first acquisition of Pseudomonas aeruginosa in cystic fibrosis. *Annals of the American Thoracic Society.* 10 (2), 108-14.

Rawat, N. & Kumar, P. (2023) Interventions for improving indoor and outdoor air quality in and around schools. *Science of the Total Environment.* 858, 159813.

Rebecca Jane, C. *et al.* (2020) Using digital technology for home monitoring, adherence and self-management in cystic fibrosis: a state-of-the-art review. *Thorax.* 75 (1), 72.

Rehan, M. & Munir, S. (2022) Analysis and Modeling of Air Pollution in Extreme Meteorological Conditions: A Case Study of Jeddah, the Kingdom of Saudi Arabia. *Toxics.* 10 (7).

Ren, J. *et al.* (2016) Approaches to prevent the patients with chronic airway diseases from exacerbation in the haze weather. *Journal of Thoracic Disease.* 8 (1), E1-7.

Robertson, S. *et al.* (2019) Bioaerosol exposure from composting facilities and health outcomes in workers and in the community: A systematic review update. *International journal of hygiene and environmental health.* 222 (3), 364-386.

Roca-Barcelo, A. *et al.* (2020) Risk of respiratory hospital admission associated with modelled concentrations of Aspergillus fumigatus from composting facilities in England. *Environmental Research.* 183, 108949.

Ronaldson, A. *et al.* (2022) Associations between air pollution and multimorbidity in the UK Biobank: A cross-sectional study. *Front Public Health.* 10, 1035415.

Rowe, S. M. *et al.* (2017) Tezacaftor–Ivacaftor in Residual-Function Heterozygotes with Cystic Fibrosis. *New England Journal of Medicine.* 377 (21), 2024-2035.

Ryall, B. *et al.* (2008) Pseudomonas aeruginosa, cyanide accumulation and lung function in CF and non-CF bronchiectasis patients. *European Respiratory Journal.* 32 (3), 740-747.

SAHSU (2024) Small Area Health Statistics Unit. https://www.imperial.ac.uk/school-public-health/epidemiology-and-biostatistics/small-area-health-statistics-unit/. [Accessed 23 May. 2019].

Salsgiver, E. L. *et al.* (2016) Changing epidemiology of the respiratory bacteriology of patients with cystic fibrosis. *Chest.* 149 (2), 390-400.

Salvatore, D. *et al.* (2019) Patients with cystic fibrosis having a residual function mutation: Data from the Italian registry. *Journal of Pediatric Pulmonology.* 54 (2), 150-157.

Samake, A. *et al.* (2017) The unexpected role of bioaerosols in the Oxidative Potential of PM. *International Journal of Scientific Reports.* 7 (1), 10978.

Samoli, E. *et al.* (2005) Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project. *Environmental Health Perspectives.* 113 (1), 88-95.

Sanders, D. B. *et al.* (2011) Pulmonary exacerbations are associated with subsequent FEV1 decline in both adults and children with cystic fibrosis. *Journal of Pediatric Pulmonology.* 46 (4), 393-400.

Sandstrom, T. *et al.* (1991) Inflammatory cell response in bronchoalveolar lavage fluid after nitrogen dioxide exposure of healthy subjects: a dose-response study. *The European respiratory journal.* 4 (3), 332-9.

Sang, S. *et al.* (2022) The global burden of disease attributable to ambient fine particulate matter in 204 countries and territories, 1990–2019: A systematic analysis of the Global Burden of Disease Study 2019. *Ecotoxicology and Environmental Safety.* 238, 113588.

Saunders, R. V. *et al.* (2016) Chronic Aspergillus fumigatus colonization of the pediatric cystic fibrosis airway is common and may be associated with a more rapid decline in lung function. *Medical Mycology.* 54 (5), 537-43.

Sawicki, G. S. *et al.* (2022) Rate of Lung Function Decline in People with Cystic Fibrosis Having a Residual Function Gene Mutation. *Pulmonary Therapy.* 8 (4), 385-395.

Sbihi, H. *et al.* (2017) Asthma Trajectories in a Population-based Birth Cohort. Impacts of Air Pollution and Greenness. *American Journal of Respiratory and Critical Care Medicine.* 195 (5), 607-613.

Schaedel, C. *et al.* (2002) Predictors of deterioration of lung function in cystic fibrosis. *Journal of Pediatric Pulmonology.* 33 (6), 483-91.

Schechter, M. S. (2011) Nongenetic influences on cystic fibrosis outcomes. *Current Opinion in Pulmonary Medicine.* 17 (6), 448-54.

Schechter, M. S. *et al.* (2001) The association of socioeconomic status with outcomes in cystic fibrosis patients in the United States. *American Journal of Respiratory and Critical Care Medicine.* 163 (6), 1331-7.

Scheers, H. *et al.* (2011) Does air pollution trigger infant mortality in Western Europe? A case-crossover study. *Environmental Health Perspectives.* 119 (7), 1017-1022.

Schraufnagel, D. E. *et al.* (2019a) Air Pollution and Noncommunicable Diseases: A Review by the Forum of International Respiratory Societies' Environmental Committee, Part 1: The Damaging Effects of Air Pollution. *Chest.* 155 (2), 409-416.

Schraufnagel, D. E. *et al.* (2019b) Air Pollution and Noncommunicable Diseases: A Review by the Forum of International Respiratory Societies' Environmental Committee, Part 2: Air Pollution and Organ Systems. *Chest.* 155 (2), 417-426.

Senthilselvan, A. *et al.* (2010) Excess longitudinal decline in lung function in grain farmers. *Journal of agromedicine.* 15 (2), 157-65.

Serdar, C. C. *et al.* (2021) Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia Medica: Casopis Hrvatskoga Drustva Medicinskih Biokemicara.* 31 (1), 010502.

Shin, S. *et al.* (2020) Association Between Road Traffic Noise and Incidence of Diabetes Mellitus and Hypertension in Toronto, Canada: A Population-Based Cohort Study. *Journal of the American Heart Association.* 9 (6), e013021.

Singh, V. K. & Schwarzenberg, S. J. (2017) Pancreatic insufficiency in Cystic Fibrosis. *Journal of Cystic Fibrosis.* 16, S70-S78.

Snyder, A. H. *et al.* (2002) Acute effects of aerosolized S-nitrosoglutathione in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine.* 165 (7), 922-926.

Southern, K. W. *et al.* (2023) Standards for the care of people with cystic fibrosis; establishing and maintaining health. *Journal of Cystic Fibrosis.*

Stanford, G. E. *et al.* (2021) Pulmonary Exacerbations in Adults With Cystic Fibrosis: A Grown-up Issue in a Changing Cystic Fibrosis Landscape. *Chest.* 159 (1), 93-102.

Stenbit, A. E. & Flume, P. A. (2011) Pulmonary exacerbations in cystic fibrosis. 17 (6), 442-447.

Stevens, D. A. *et al.* (2003) Allergic bronchopulmonary aspergillosis in cystic fibrosis--state of the art: Cystic Fibrosis Foundation Consensus Conference. *Clinical Infectious Diseases.* 37 Suppl 3, S225-64.

Strassmann, A. *et al.* (2021) NO2 and PM2.5 Exposures and Lung Function in Swiss Adults: Estimated Effects of Short-Term Exposures and Long-Term Exposures with and without Adjustment for Short-Term Deviations. *Environmental Health Perspectives.* 129 (1), 17009.

Sun, J. *et al.* (2020) The effect of ambient fine particulate matter (PM2.5) on respiratory diseases in China: a systematic review and meta-analysis. *Stochastic Environmental Research and Risk Assessment.* 34 (3), 593-610.

Szczesniak, R. *et al.* (2017a) Use of FEV1 in cystic fibrosis epidemiologic studies and clinical trials: A statistical perspective for the clinical researcher. *Journal of Cystic Fibrosis.* 16 (3), 318-326.

Szczesniak, R. *et al.* (2020) Influences of environmental exposures on individuals living with cystic fibrosis. *Expert Review of Respiratory Medicine.* 14 (7), 737-748.

Szczesniak, R. D. *et al.* (2017b) Phenotypes of Rapid Cystic Fibrosis Lung Disease Progression during Adolescence and Young Adulthood. *American Journal of Respiratory and Critical Care Medicine.* 196 (4), 471-478.

Szczesniak, R. D. *et al.* (2013) A semiparametric approach to estimate rapid lung function decline in cystic fibrosis. *Annals of Epidemiology.* 23 (12), 771-7.

Taha, M. P. M. *et al.* (2006) Bioaerosol releases from compost facilities: Evaluating passive and active source terms at a green waste facility for improved risk assessments. *Atmospheric Environment.* 40 (6), 1159-1169.

Takahashi, A. *et al.* (2006) Mitogenic signalling and the p16INK4a-Rb pathway cooperate to enforce irreversible cellular senescence. *Nature Cell Biology.* 8 (11), 1291-7.

Takebayashi, T. *et al.* (2022) Exposure to PM(2.5) and Lung Function Growth in Pre- and Early-Adolescent Schoolchildren: A Longitudinal Study Involving Repeated Lung Function Measurements in Japan. *Annals of the American Thoracic Society.* 19 (5), 763-772.

Talaminos Barroso, A. *et al.* (2018) Factors Affecting Lung Function: A Review of the Literature. *Archivos de Bronconeumologia.* 54 (6), 327-332.

Taylor-Robinson, D. *et al.* (2018) Data Resource Profile: The UK Cystic Fibrosis Registry. *International Journal of Epidemiology.* 47 (1), 9-10e.

Taylor-Robinson, D. *et al.* (2015) Health Trajectories in People with Cystic Fibrosis in the UK: Exploring the Effect of Social Deprivation. In: Burton-Jeangros, C., Cullati, S., Sacker, A. & Blane, D. (eds.) *A Life Course Perspective on Health Trajectories and Transitions*. Cham (CH), pp. 85-110.

Taylor-Robinson, D. & Schechter, M. S. (2011) Health inequalities and cystic fibrosis. *British Medical Journal.* 343, d4818.

Taylor-Robinson, D. *et al.* (2013) The effect of social deprivation on clinical outcomes and the use of treatments in the UK cystic fibrosis population: a longitudinal study. *The Lancet Respiratory Medicine.* 1 (2), 121-8.

Taylor-Robinson, D. *et al.* (2012) Understanding the natural progression in %FEV1 decline in patients with cystic fibrosis: a longitudinal study. *Thorax.* 67 (10), 860-6.

Taylor-Robinson, D. C. *et al.* (2014) Low socioeconomic status is associated with worse lung function in the Danish cystic fibrosis population. *The European respiratory journal.* 44 (5), 1363-6.

The Lancet (2006) WHO's global air-quality guidelines. *The Lancet.* 368 (9544), 1302.

The World Bank. (2018) *WHAT A WASTE:Trends in Solid Waste Management,* https://datatopics.worldbank.org/what-a-waste/trends_in_solid_waste_management.html [Accessed 23rd Sep 2020].

Theodore Lytras. (2022) *Package 'rspiro'.* https://cran.r-project.org/web/packages/rspiro/rspiro.pdf.

Tian, L. *et al.* (2019) Spatial patterns and effects of air pollution and meteorological factors on hospitalization for chronic lung diseases in Beijing, China. *Science China Life Sciences.* 62 (10), 1381-1388.

Tierney. et. al. (2022) "The R Journal: brolgar: An R package to BRowse Over Longitudinal Data Graphically and Analytically in R". *The R Journal.*

Tiotiu, A. I. *et al.* (2020) Impact of Air Pollution on Asthma Outcomes. *International Journal of Environmental Research and Public Health.* 17 (17).

Townsend, E. A. *et al.* (2012) Sex differences and sex steroids in lung health and disease. *Endocrine Reviews.* 33 (1), 1-47.

Trasande, L. *et al.* (2016) Particulate matter exposure and preterm birth: estimates of US attributable burden and economic costs. *Environmental Health Perspectives.* 124 (12), 1913-1918.

Tsai, S.-S. *et al.* (2013) Fine particulate air pollution and hospital admissions for chronic obstructive pulmonary disease: a case-crossover study in Taipei. *International Journal of Environmental Research and Public Health.* 10 (11), 6015-6026.

Tseng, C.-M. *et al.* (2013) The effect of cold temperature on increased exacerbation of chronic obstructive pulmonary disease: a nationwide study. *PloS One.* 8 (3), e57066.

UK CF Registry. (2020) UK Cystic Fibrosis Registry Annual Data Report 2020. https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources [Accessed 20th Feb 2022].

UK CF Registry. (2021) UK Cystic Fibrosis Registry Annual Data Report 2021. https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources [Accessed 16th June 2023].

Van Goor, F. *et al.* (2014) Effect of ivacaftor on CFTR forms with missense mutations associated with defects in protein processing or function. *Journal of Cystic Fibrosis.* 13 (1), 29-36.

van Kampen, V. *et al.* (2016) Effects of bioaerosol exposure on respiratory health in compost workers: a 13-year follow-up study. *Occupational & Environmental Medicine* 73 (12), 829-837.

VanderWeele, T. J. & Robins, J. M. (2007) Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol.* 166 (9), 1096-104.

Veit, G. *et al.* (2016) From CFTR biology toward combinatorial pharmacotherapy: expanded classification of cystic fibrosis mutations. *Molecular Biology of the Cell.* 27 (3), 424-33.

Verhoeven, J. I. *et al.* (2021) Ambient air pollution and the risk of ischaemic and haemorrhagic stroke. *The Lancet Planetary Health.* 5 (8), e542-e552.

Verma, N. *et al.* (2005) Is There Still a Gender Gap in Cystic Fibrosis? *Chest.* 128 (4), 2824-2834.

Vidal, D. G. *et al.* (2022) Chapter 6 - The influence of social and economic environment on health. In: Prata, J.C., Ribeiro, A.I. & Rocha-Santos, T. (eds.) *One Health*, Academic Press, pp. 205-229.

Vijayan, V. K. *et al.* (2015) Enhancing indoor air quality -The air filter advantage. *Lung India.* 32 (5), 473-9.

Vogel M. (2022.) *Vogel M. Childsds R package version 0.8.0 2022: Data and Methods Around Reference Values in Pediatrics* https://cran.r-project.org/web/packages/childsds/index.html [Accessed 10th March 2022].

Wainwright, C. E. *et al.* (2015) Lumacaftor–Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. *New England Journal of Medicine.* 373 (3), 220-231.

Walser, S. M. *et al.* (2015) Evaluation of exposure–response relationships for health effects of microbial bioaerosols – A systematic review. *International journal of hygiene and environmental health.* 218 (7), 577-589.

Walters, S. *et al.* (1994) Hospital care for adults with cystic fibrosis: an overview and comparison between special cystic fibrosis clinics and general clinics using a patient questionnaire. *Thorax.* 49 (4), 300-306.

Wang, L. *et al.* (2022a) Air pollution and risk of chronic obstructed pulmonary disease: The modifying effect of genetic susceptibility and lifestyle. *EBioMedicine.* 79, 103994.

Wang, W. *et al.* (2022b) Predicting daily concentrations of nitrogen dioxide, particulate matter and ozone at fine spatial scale in Great Britain. *Atmospheric Pollution Research.* 13 (8), 101506.

Wang, Y. *et al.* (2017) Long-term Exposure to PM2.5 and Mortality Among Older Adults in the Southeastern US. *Epidemiology (Cambridge, Mass.).* 28 (2), 207-214.

Wang, Y. *et al.* (2023) ADMS simulation and influencing factors of bioaerosol diffusion from BRT under different aeration modes in six wastewater treatment plants. *Water Research.* 231, 119624.

Warrier, A. & Sood, A. (2021) Home-Based Physiological Monitoring of Patients with COVID-19. *Southwest Journal of Pulmonary and Critical Care.* 23 (3), 76-88.

Wei, M. *et al.* (2020) Pollution characteristics of bioaerosols in PM 2.5 during the winter heating season in a coastal city of northern China. *Environmental Science and Pollution Research.* 27, 27750-27761.

Wellenius, G. A. *et al.* (2005) Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries. *Stroke.* 36 (12), 2549-53.

Wéry, N. (2014) Bioaerosols from composting facilities--a review. *Frontiers in Cellular and Infection Microbiology.* 4, 42.

Whitehouse, A. L. *et al.* (2020) Airway dendritic cell maturation in children exposed to air pollution. *PloS One.* 15 (5), e0232040.

WHO. (2016) World Health Organisation. Ambient air pollution: a global assessment of exposure and burden of disease. *Clean Air Journal.* 26 (2), 6.

WHO. (2022) *WHO Global Ambient Air Quality Database.* https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health [Accessed 17 June. 2023].

WHO. (2024) *Ambient air pollution.* https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/ambient-air-pollution#:~:text=Globally%2C%20air%20pollution%20it%20is%20estimated%20to%20cause,low-%20and%20middle-income%20countries%20disproportionately%20experience%20this%20burden [Accessed 14th December 2022].

Williams, B. *et al.* (2019) Estimating Aspergillus fumigatus exposure from outdoor composting activities in England between 2005 and 14. *Waste Manag.* 84, 235-244.

Williams, C. *et al.* (2016) Pathogenesis of Fungal Infections in Cystic Fibrosis. *Current fungal infection reports.* 10 (4), 163-169.

Wong, T. W. *et al.* (1999) Air pollution and hospital admissions for respiratory and cardiovascular diseases in Hong Kong. *Journal of Occupational and Environmental Medicine.* 56 (10), 679-83.

Xiao, Q. *et al.* (2017) Full-coverage high-resolution daily PM2.5 estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote Sensing of Environment.* 199, 437-446.

Xie, S. *et al.* (2021) Feasibility and acceptability of monitoring personal air pollution exposure with sensors for asthma self-management. *Asthma Research and Practice.* 7 (1), 13.

Xing, Y. F. *et al.* (2016) The impact of PM2.5 on the human respiratory system. *Journal of Thoracic Disease.* 8 (1), E69-74.

Xu, Z. *et al.* (2020) Interleukins 6/8 and cyclooxygenase-2 release and expressions are regulated by oxidative stress-JAK2/STAT3 signaling pathway in human bronchial epithelial cells exposed to particulate matter≤ 2.5 μm. *Journal of Applied Toxicology.* 40 (9), 1210-1218.

Yamada, I. (2016) Thiessen Polygons. In International Encyclopedia of Geography: People, the Earth, Environment and Technology (eds D. Richardson, N. Castree, M.F. Goodchild, A. Kobayashi, W. Liu and R.A. Marston). https://doi.org/10.1002/9781118786352.wbieg0157. pp. 1-6.

Yang, J. *et al.* (2021a) Investigating the impact of air pollution on AMI and COPD hospital admissions in the coastal city of Qingdao, China. *Frontiers of Environmental Science & Engineering.* 16 (5), 56.

Yang, X. *et al.* (2021b) Global burden of COPD attributable to ambient PM2. 5 in 204 countries and territories, 1990 to 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Science of the Total Environment.* 796, 148819.

Yang, X. Y. *et al.* (2020) Acute Effects of Individual Exposure to Fine Particulate Matter on Pulmonary Function in Schoolchildren. *Biomedical and Environmental Sciences.* 33 (9), 647-659.

Yoo, E. H. *et al.* (2021) The Impact of Individual Mobility on Long-Term Exposure to Ambient PM(2.5): Assessing Effect Modification by Travel Patterns and Spatial Variability of PM(2.5). *International Journal of Environmental Research and Public Health.* 18 (4).

Yu, H. *et al.* (2012) Ivacaftor potentiation of multiple CFTR channels with gating mutations. *Journal of Cystic Fibrosis.* 11 (3), 237-245.

Yuan, Y. *et al.* (2023) Excess mortality associated with high ozone exposure: A national cohort study in China. *Environmental Science and Ecotechnology.* 15, 100241.

Zemel, B. S. *et al.* (2000) Longitudinal relationship among growth, nutritional status, and pulmonary function in children with cystic fibrosis: analysis of the Cystic Fibrosis Foundation National CF Patient Registry. *The Journal of pediatrics.* 137 (3), 374-80.

Zhang, W. *et al.* (2022) The relationship between particulate matter and lung function of children: A systematic review and meta-analysis. *Environmental Pollution.* 309, 119735.

Zhang, Y. *et al.* (2023) Effect of Acute PM2.5 Exposure on Lung Function in Children: A Systematic Review and Meta-Analysis. *Journal of Asthma and Allergy.* 16, 529-540.

Zhao, Q. *et al.* (2021) Air pollution during infancy and lung function development into adolescence: The GINIplus/LISA birth cohorts study. *Environment International.* 146, 106195.

Zheng, X. Y. *et al.* (2021) Short-term exposure to ozone, nitrogen dioxide, and sulphur dioxide and emergency department visits and hospital admissions due to asthma: A systematic review and meta-analysis. *Environment International.* 150, 106435.

Zielenski, J. (2000) Genotype and phenotype in cystic fibrosis. *Respiration.* 67 (2), 117-33.

7 Appendices

7.1 Chapter II: Boxplots of the relationship between disease severity (ppFEV$_1$ and PEx) and air pollutants (NO_2 and $PM_{2.5)}$ in pwCF in GB

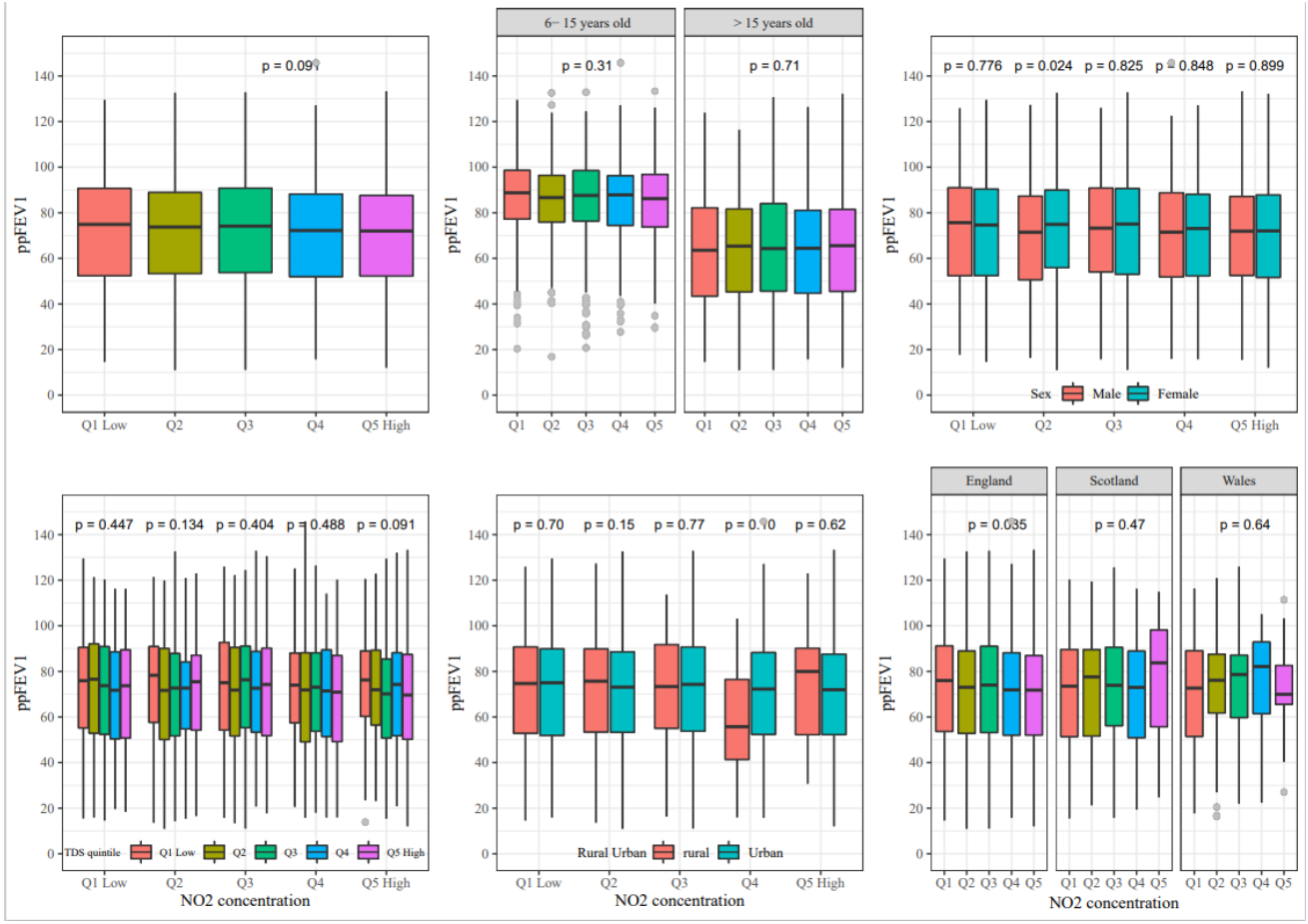


Figure 7.1: Boxplot of the relationship between ppFEV$_1$ and NO_2 by covariates in pwCF in GB

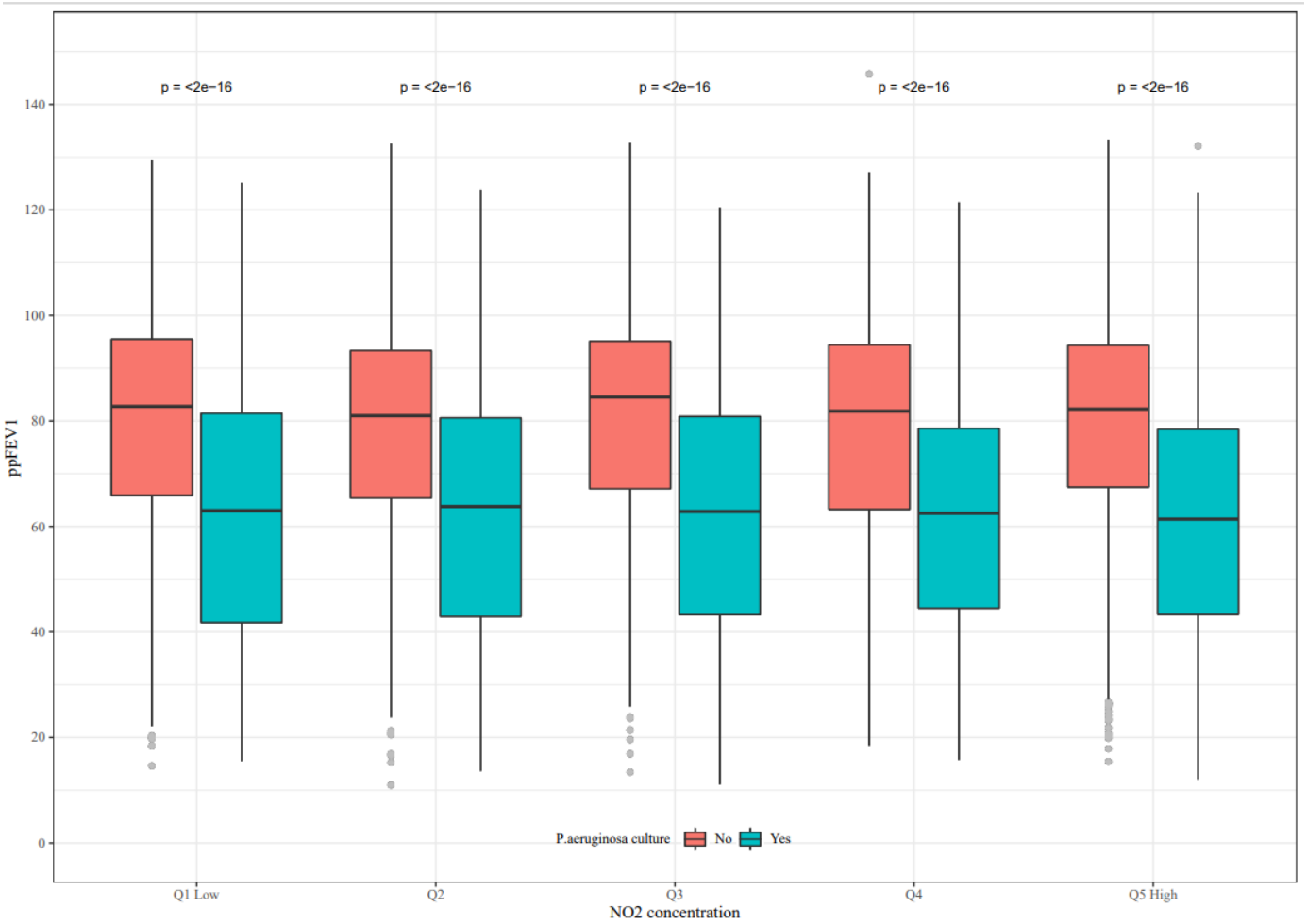


Figure 7.2: Boxplot of the relationship between ppFEV$_1$ and NO_2 by *P. aeruginosa* in pwCF in GB

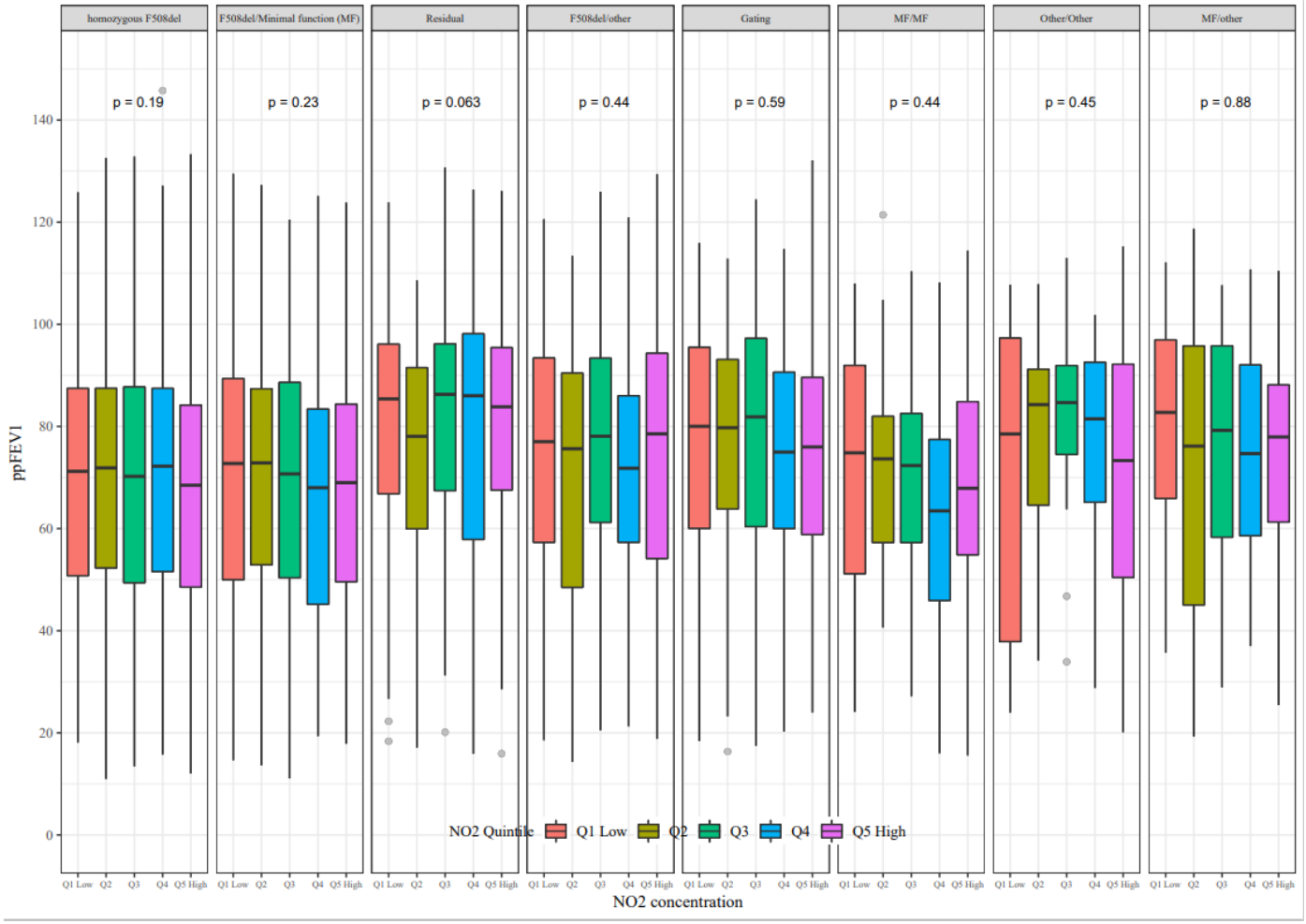


Figure 7.3: Boxplot of the relationship between $ppFEV_1$ and NO_2 by genotype in pwCF in GB

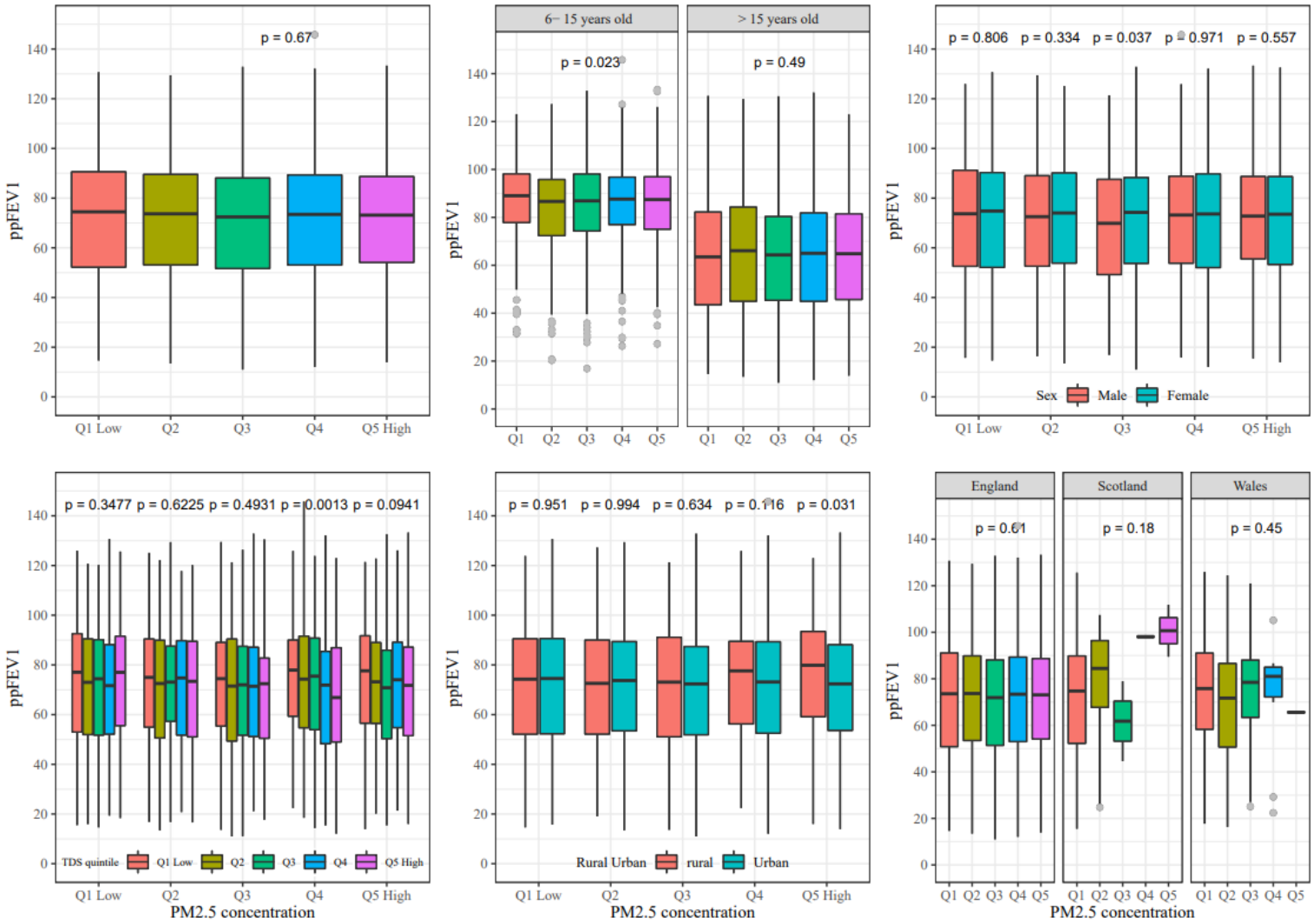


Figure 7.4: Boxplot of the relationship between ppFEV$_1$ and PM$_{2.5}$ by covariates in pwCF in GB

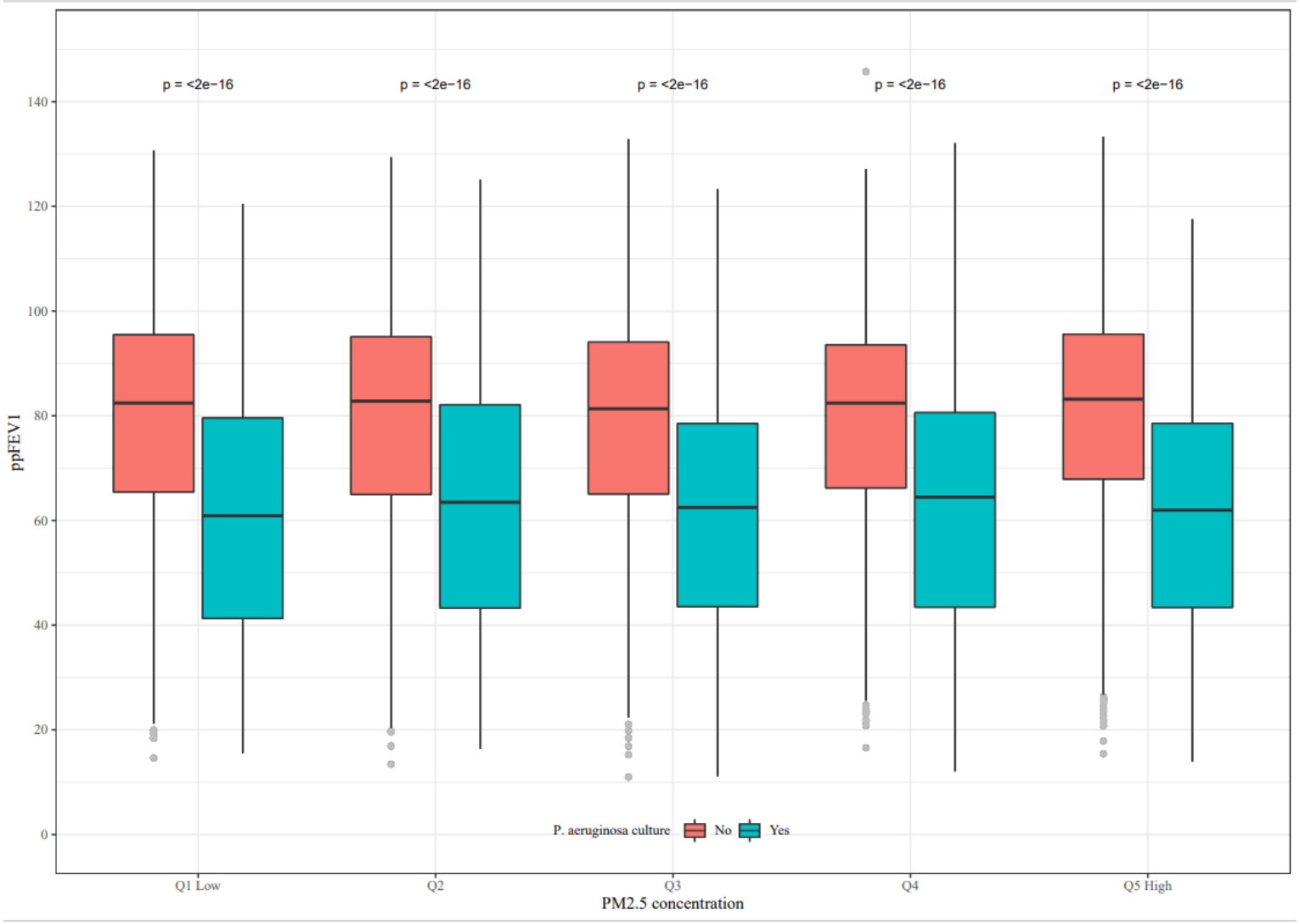


Figure 7.5: Boxplot of the relationship between $ppFEV_1$ and $PM_{2.5}$ by *P. aeruginosa* in pwCF in GB

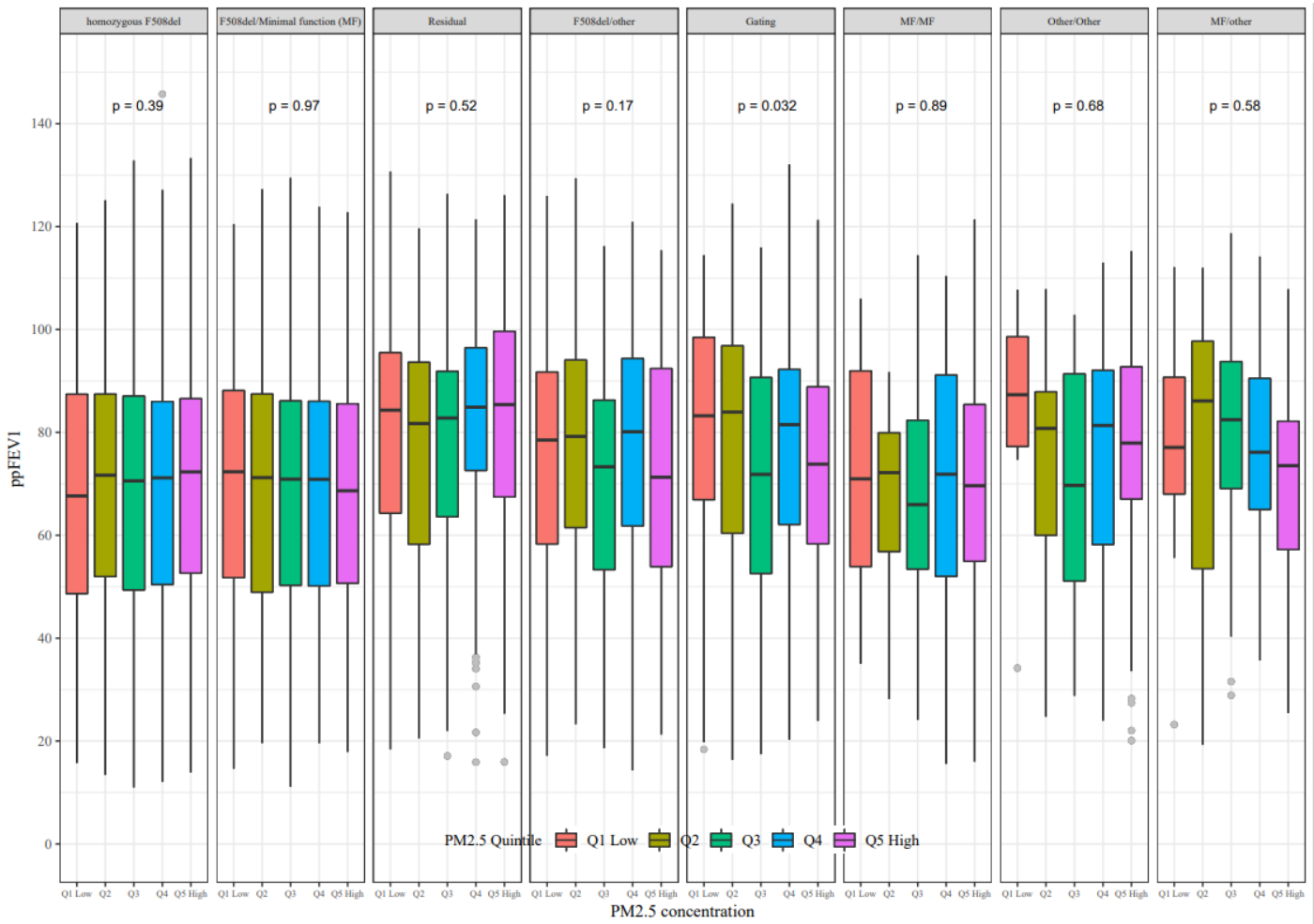


Figure 7.6: Boxplot of the relationship between ppFEV$_1$ and PM$_{2.5}$ by genotype in pwCF in GB

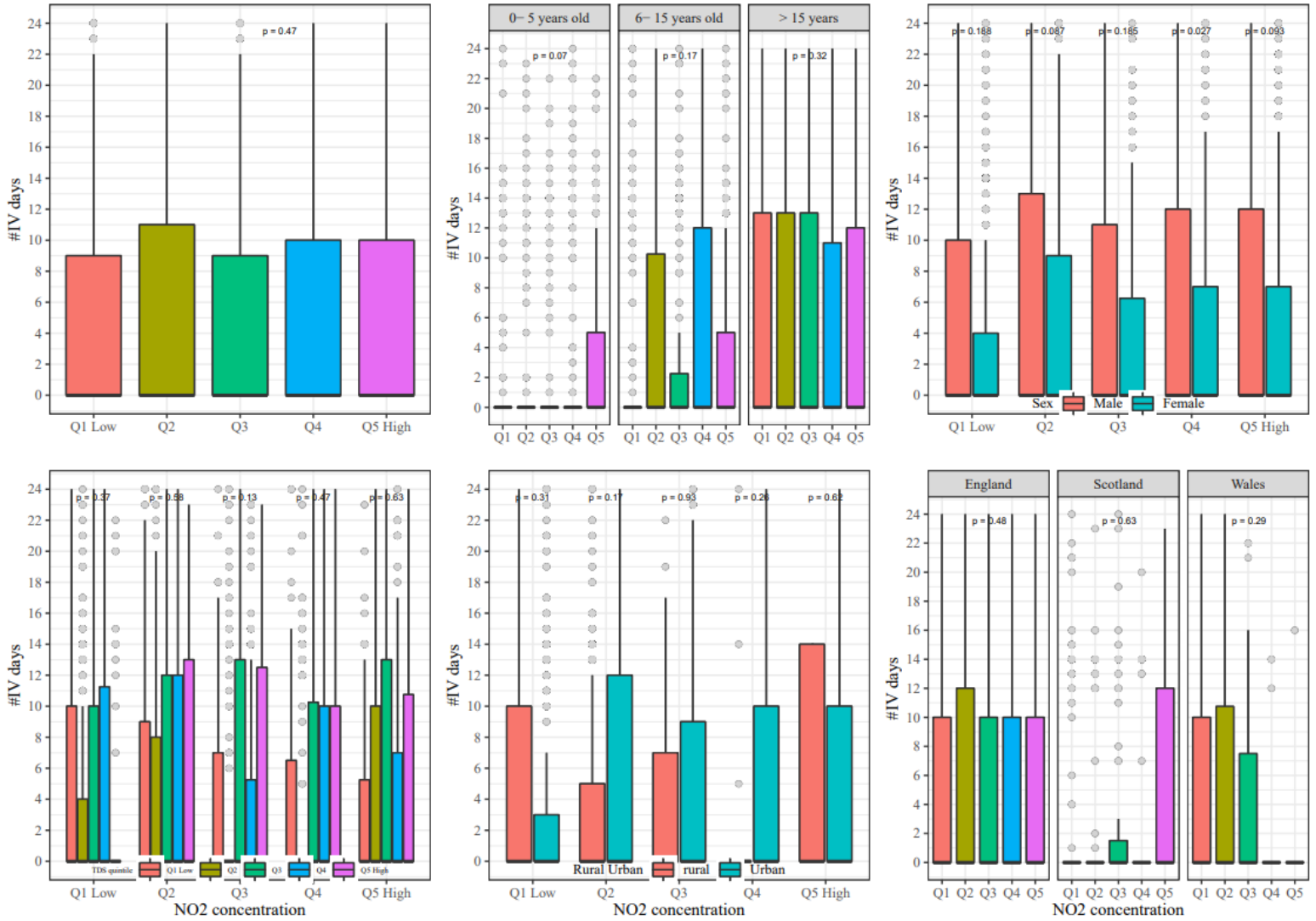


Figure 7.7: Boxplot of the relationship between #IVdays and NO_2 by covariates in pwCF in GB

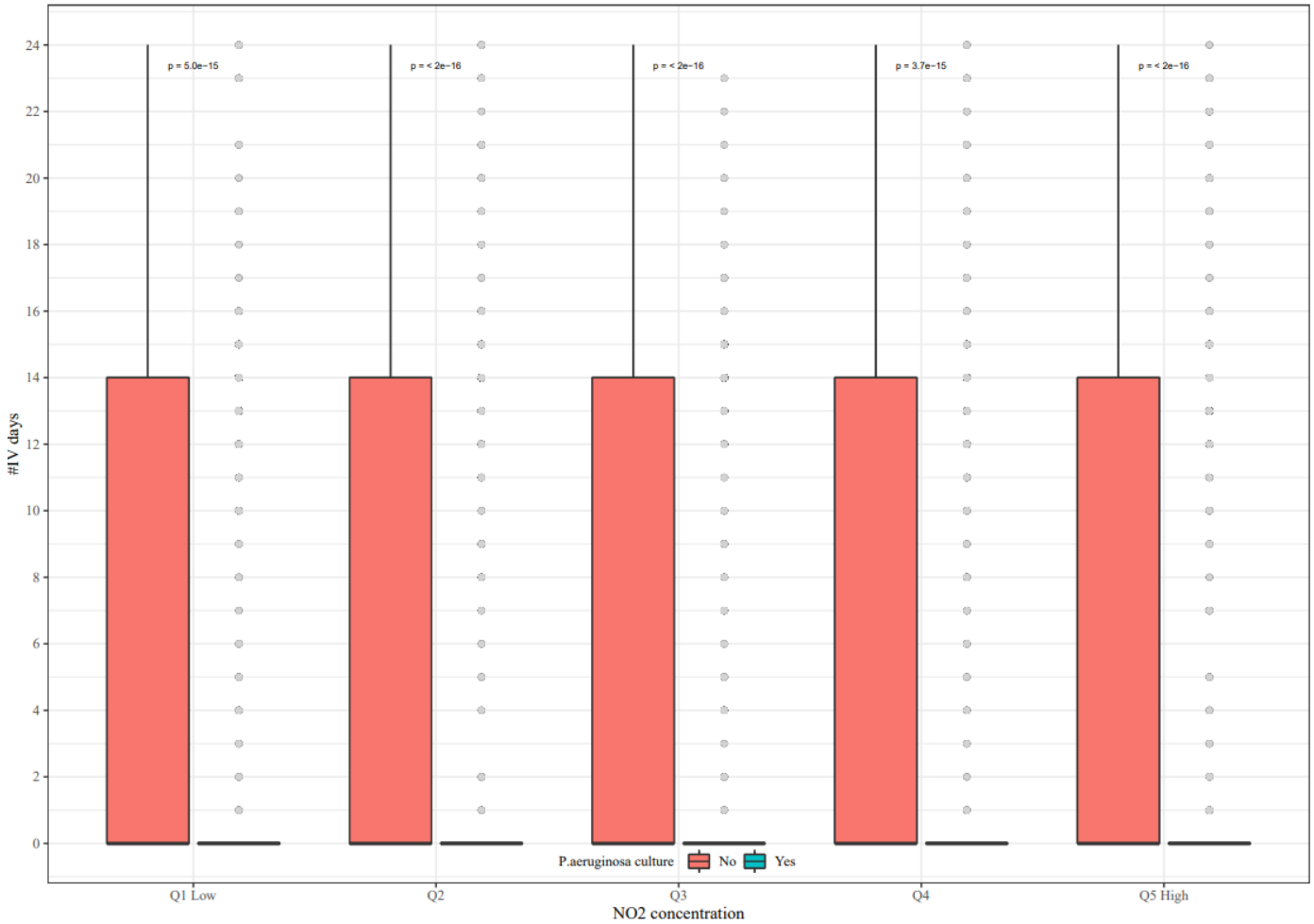


Figure 7.8: Boxplot of the relationship between #IVdays and NO_2 by *P. aeruginosa* in pwCF in GB

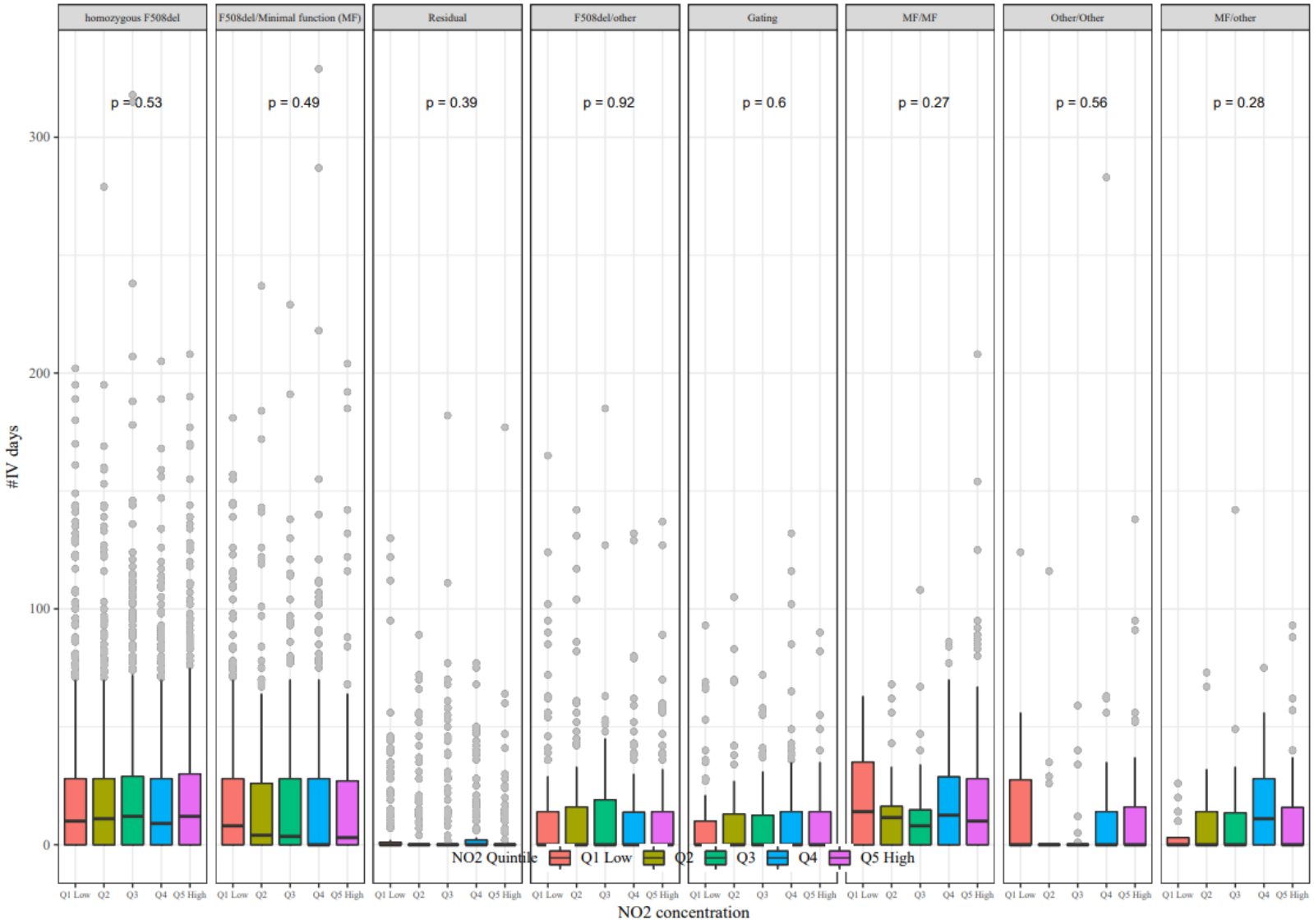


Figure 7.9: Boxplot of the relationship between #IVdays and NO_2 by genotype in pwCF in GB

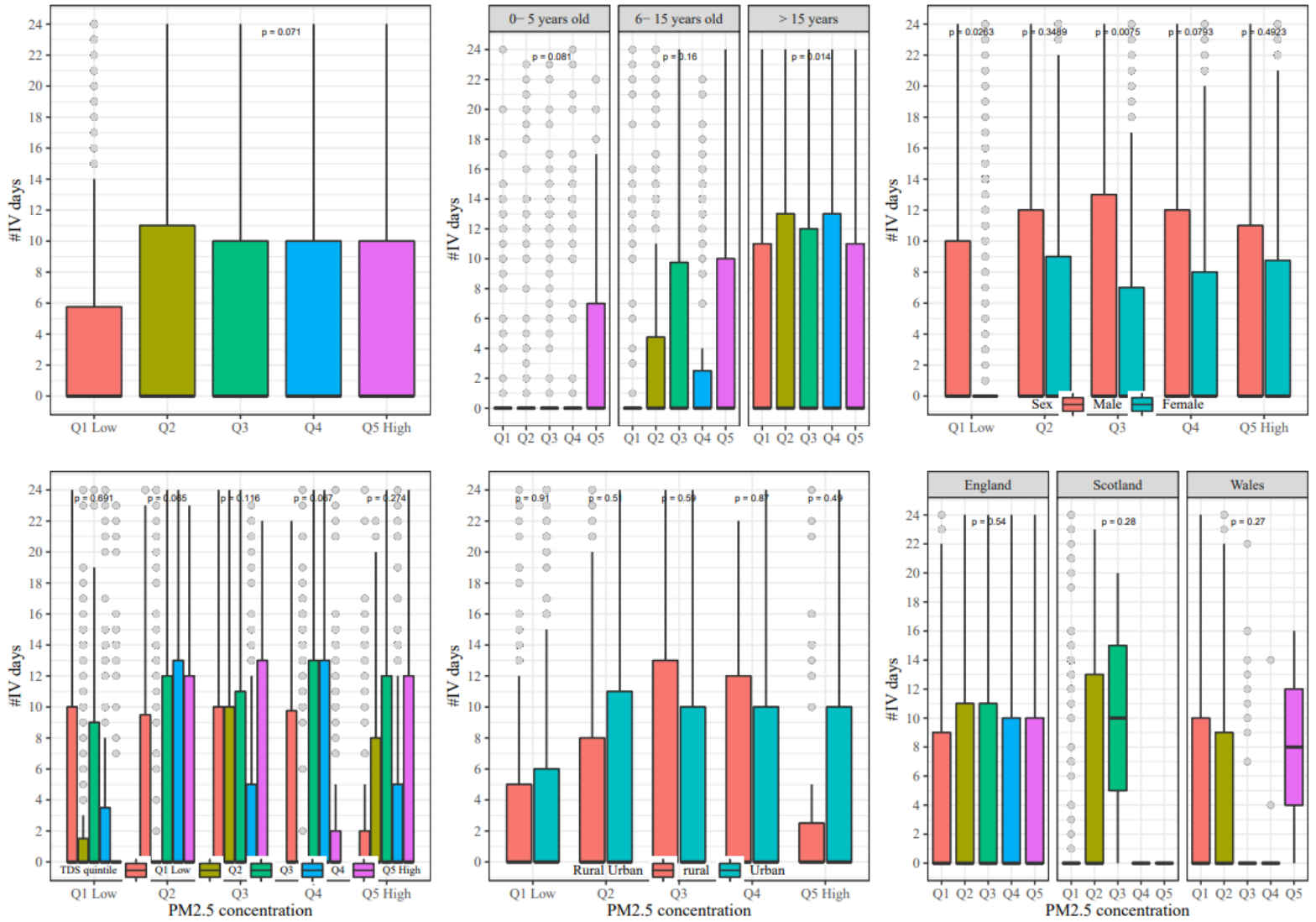


Figure 7.10: Boxplot of the relationship between #IVdays and $PM_{2.5}$ by confounders in pwCF in GB

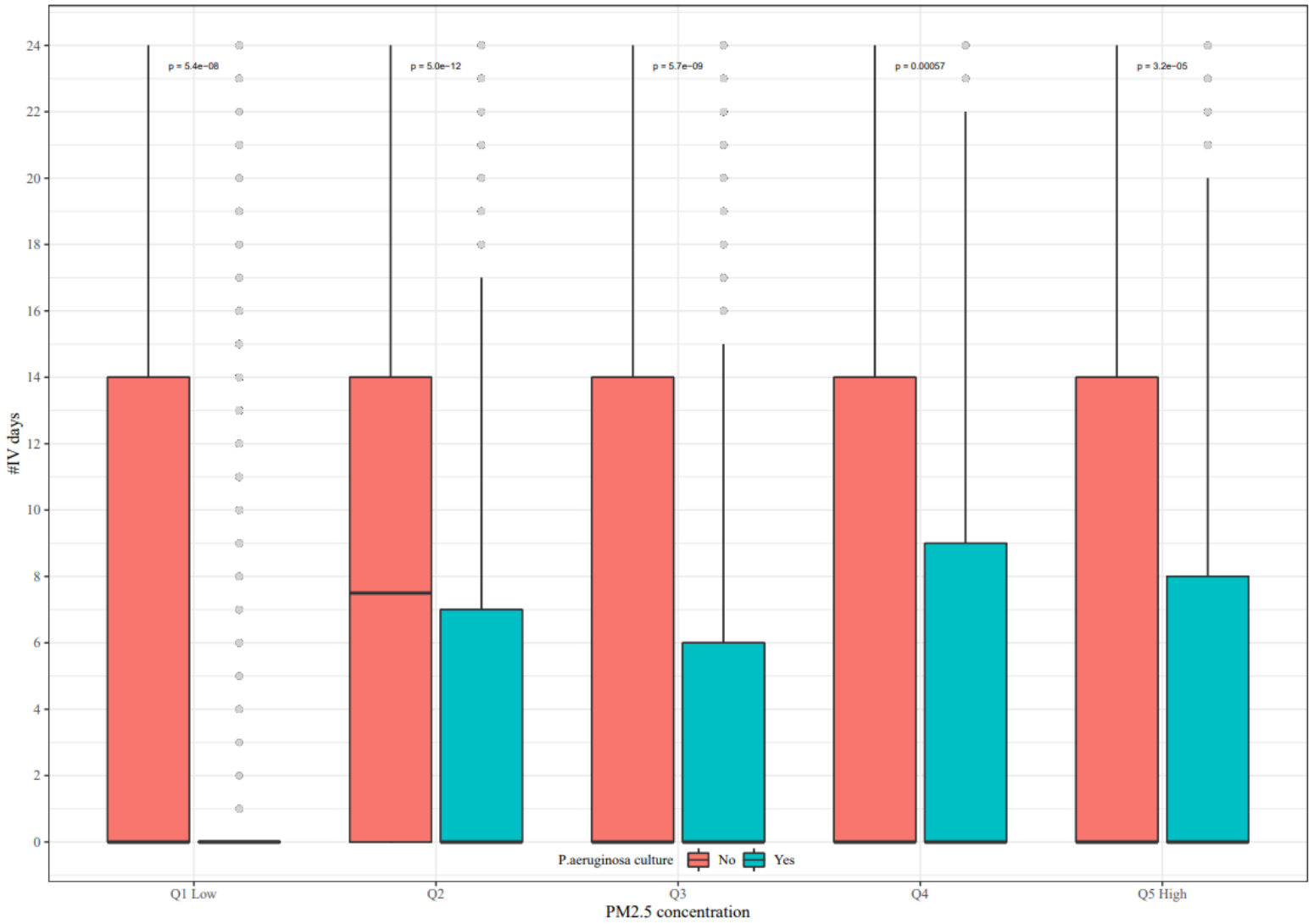


Figure 7.11: Boxplot of the relationship between #IVdays and $PM_{2.5}$ by *P. aeruginosa* in pwCF in GB

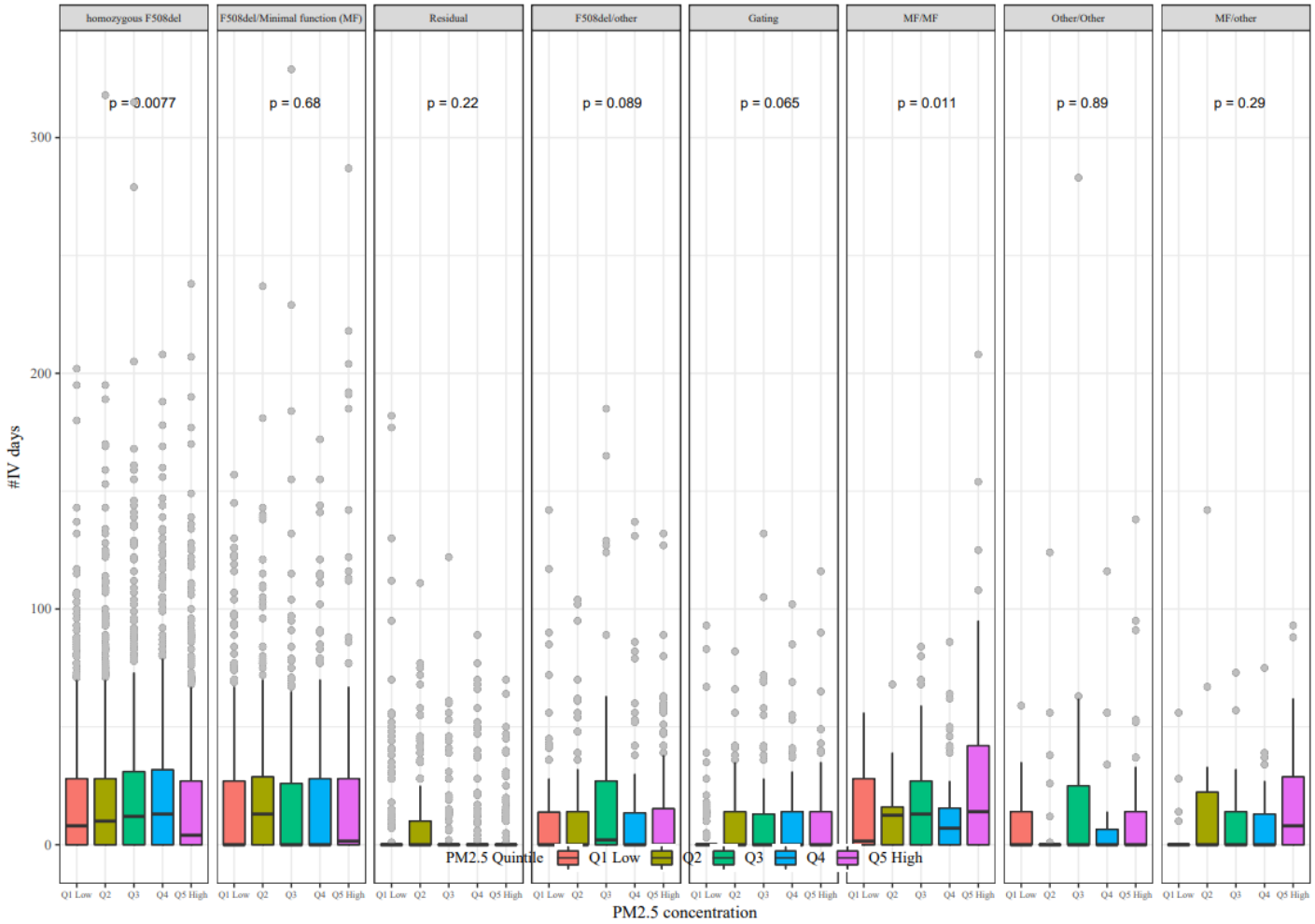


Figure 7.12: Boxplot of the relationship between #IVdays and $PM_{2.5}$ by genotype in pwCF in GB

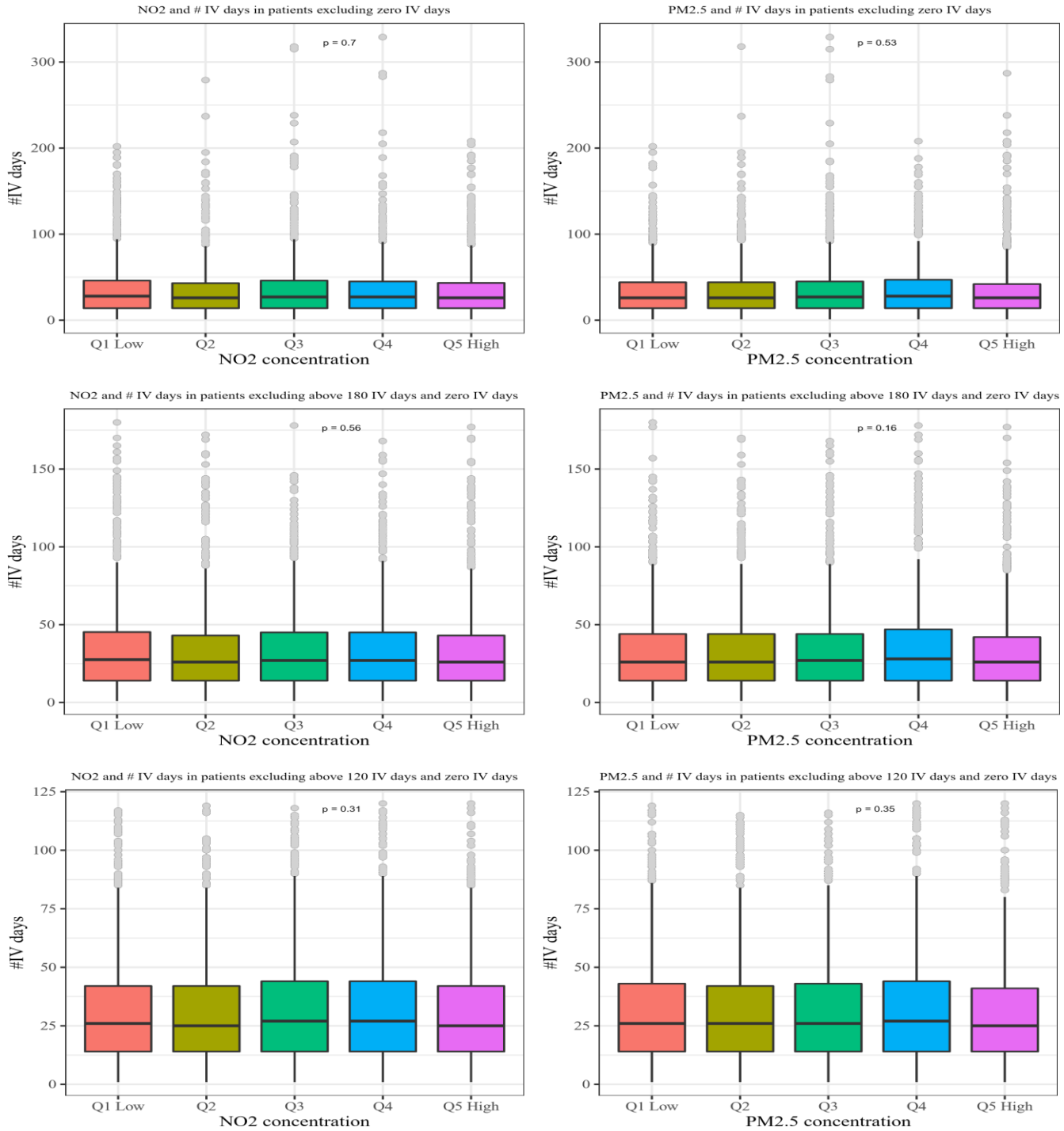


Figure 7.13: Boxplot of the relationship between subgroup of #IVdays and air pollutants (NO_2 and $PM_{2.5}$) in pwCF in GB

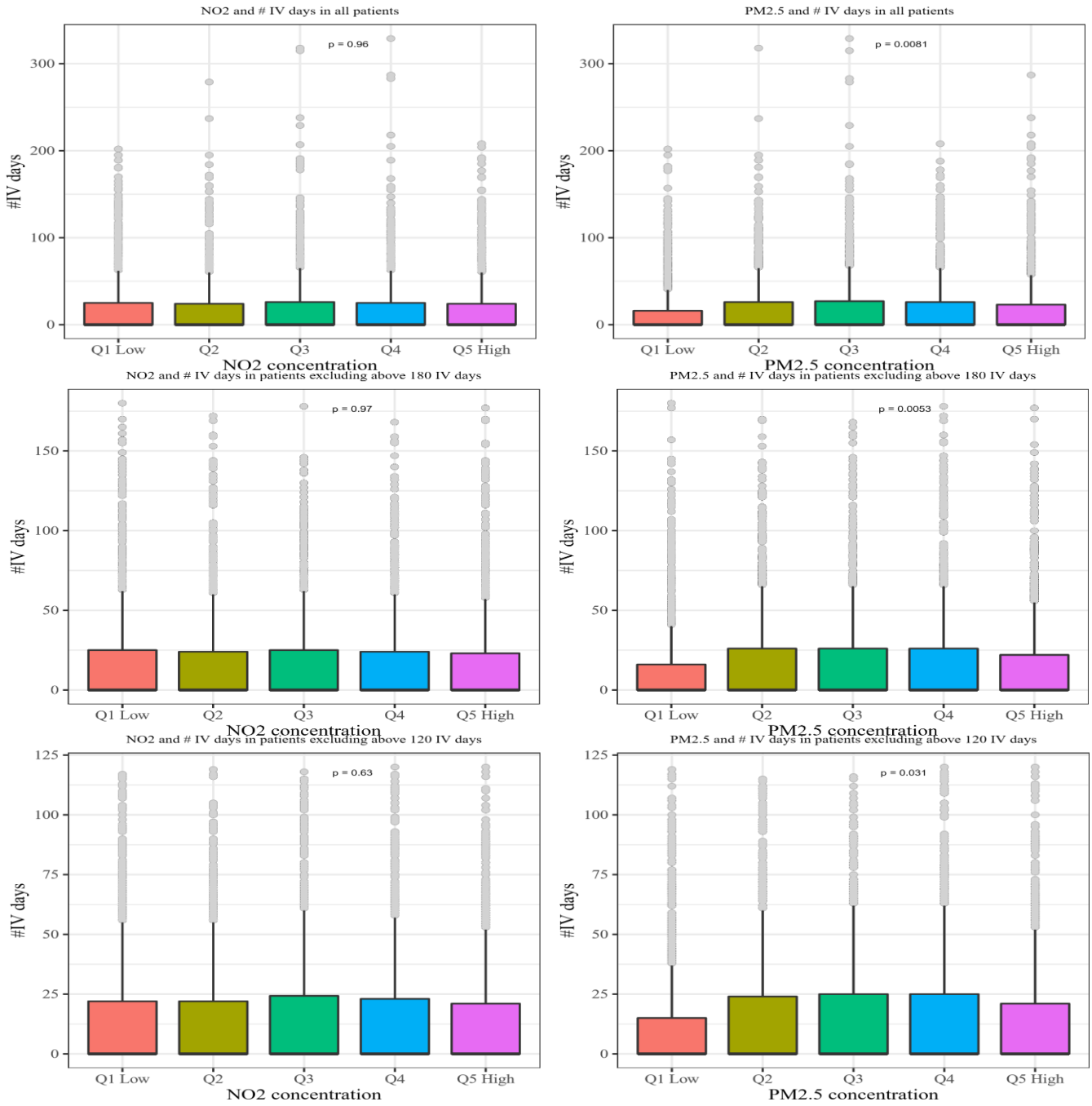


Figure 7.14: Boxplot of the relationship between subgroup of #IVdays and air pollutants (NO_2 and $PM_{2.5}$) in pwCF in GB

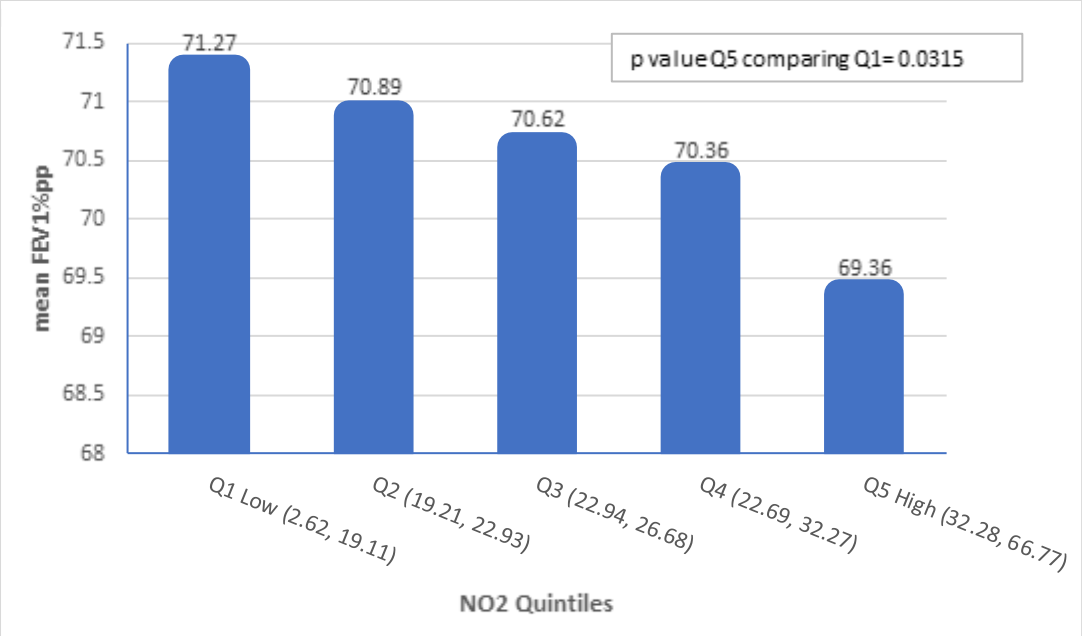


Figure 7.15 : ppFEV$_1$ by quintile of NO_2 Concentrations in GB

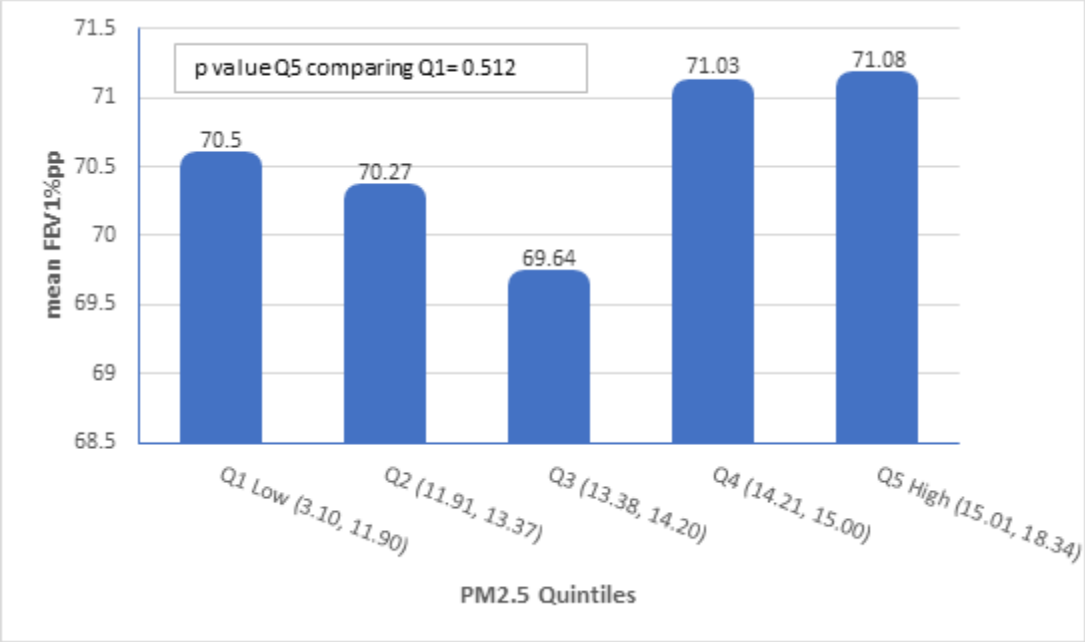


Figure 7.16: ppFEV$_1$ by quintile of PM$_{2.5}$ Concentrations in GB

Table 7.1: Linear regressions between air pollutants and lung function in children (age= 6 – 15 years) with CF in the UK CF Registry in GB

	$ppFEV_1$											
	All children (age 6 – 15 years) n = 2,268						**Children with CF who didn't change their residential postcode between 2012 and 2016 n = 1,448**					
	Unadjusted			**Adjusted**			**Unadjusted**			**Adjusted**		
Characteristic	**β**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**	**B**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (13.52, 17.02)	-1.33	-3.55, 0.88	0.239	-0.64	-2.69, 1.42	0.542	-1.72	-4.49, 1.05	0.223	-0.07	-2.64, 2.50	0.959
Q3 (17.03, 19.88)	-0.63	-2.85, 1.59	0.578	0.12	-1.98, 2.23	0.907	-1.08	-3.83, 1.66	0.437	0.28	-2.31, 2.88	0.831
Q4 (19.89, 23.67)	-2.18	-4.45, 0.09	0.060	-0.98	-3.19, 1.22	0.381	-3.08	-5.89, -0.27	**0.032**	-1.52	-4.24, 1.19	0.272
Q5 – High (23.68, 71.18)	-1.88	-4.19, 0.43	0.111	-0.70	-3.05, 1.64	0.556	-2.41	-5.27, 0.46	0.100	-0.88	-3.76, 2.00	0.549
AIC		18,481.5			18,097.4			11,749.4			11,509.4	
$PM_{2.5}$ (2015 model) (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (8.56, 9.41)	-3.76	-6.03, -1.50	0.001	-3.02	- 0.94, 2.70	0.342	-2.96	-5.79, -0.12	0.041	-1.70	- 4.35, 0.86	0.190
Q3 (9.42, 9.87)	-2.49	-4.79, -0.19	0.034	-1.56	-1.36, 2.30	0.616	-2.15	-5.02, 0.72	0.141	-1.16	-3.79, 1.47	0.388
Q4 (9.88, 10.27)	-1.50	-3.79, 0.78	0.197	-0.76	-0.82, 2.85	0.277	-2.01	-4.82, 0.79	0.160	-1.36	-3.94, 1.22	0.302
Q5 – High (10.28, 13.87)	-1.73	-4.03, 0.58,	0.141	-0.33	0.70, 4.40	**0.007**	-1.32	-4.19, 1.56	0.370	-0.22	-2.89, 2.45	0.870
AIC		18,475.3			18,090.5			11,750.7			11,509.3	

CF = cystic fibrosis; $ppFEV_1$ = percent predicted FEV_1; AIC= Akaike information criterion; GB= Great Britain
NO_2 model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation, country and sex
$PM_{2.5}$ model adjusted for BMI percentile, age, *P. aeruginosa,* genotype, deprivation and sex

Table 7.2: Linear regression between air pollutants (NO_2 and $PM_{2.5}$) and lung function in adults (age ≥ 16) with CF in the UK CF Registry in GB who didn't change their residential postcode between 2012 and 2016

	ppFEV$_1$											
	All adults (age ≥ 16) **n = 5,203**						**Adults with CF who didn't change their residential postcode between 2012 and 2016** **n = 2,891**					
	Unadjusted			**Adjusted**			**Unadjusted**			**Adjusted**		
Characteristic	**β**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**	**B**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (13.52, 17.02)	0.52	-1.59, 2.62	0.631	0.34	-1.53, 2.21	0.722	0.01	-2.84, 2.87	0.992	-0.09	-2.61, 2.42	0.942
Q3 (17.03, 19.88)	1.32	-0.79, 3.43	0.219	1.49	-0.41, 3.38	0.124	0.06	-2.84, 2.97	0.965	0.74	-1.85, 3.34	0.574
Q4 (19.89, 23.67)	-0.04	-2.15, 2.06	0.965	1.38	-0.57, 3.32	0.165	-1.41	-4.29, 1.48	0.339	0.19	-2.46, 2.85	0.887
Q5 – High (23.68, 71.18)	0.38	-1.71, 2.46	0.724	2.11	0.07, 4.16	**0.043**	-1.03	--3.93, 1.87	0.487	1.08	-1.76, 3.92	0.454
AIC		45657.0			44304.1			25456.4			24681.7	
$PM_{2.5}$ (2015 model) (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (8.56, 9.41)	1.34	-0.74, 3.43	0.208	0.88	- 0.94, 2.70	0.342	0.99	-1.85, 3.83	0.494	0.69	-1.78, 3.17	0.581
Q3 (9.42, 9.87)	-0.54	-2.63, 1.55	0.614	0.47	-1.36, 2.30	0.616	-1.84	-4.68, 0.99	0.202	-0.28	-2.76, 2.19	0.823
Q4 (9.88, 10.27)	0.01	-2.07, 2.10	0.990	1.02	-0.82, 2.85	0.277	-0.65	-3.46, 2.16	0.650	0.84	-1.66, 3.27	0.522
Q5 – High (10.28, 13.87)	0.35	-1.72, 2.42	0.740	2.54	0.70, 4.40	**0.007**	0.34	-2.51, 3.19	0.815	2.83	0.31, 5.35	**0.028**
AIC		45655.7			44302.2			25453.9			34678.5	

CF = cystic fibrosis; ppFEV$_1$ = percent predicted FEV$_1$; AIC= Akaike information criterion; GB= Great Britain

NO_2 model adjusted for *P. aeruginosa*, absolute BMI, age, genotype, deprivation, country and sex (double check country in adults)

$PM_{2.5}$ model adjusted for *P. aeruginosa*, absolute BMI, age, genotype, deprivation and sex

Table 7.3: Zero-inflated negative binomial regression analysis between NO_2 and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB

	#IVdays											
	Count model coefficient						Zero-inflation model coefficient					
	Unadjusted			Adjusted*			Unadjusted			Adjusted*		
Characteristic	**exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Ex(β)**	**95% CI**	**p-value**
All children (0-15 years) (n= 3,762)												
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref	-			-		Ref	-			-	
Q2 (13.52, 17.02)	1.01	0.89, 1.14	0.891	1.04	0.92, 1.18	0.506	0.91	0.74, 1.12	0.367	0.90	0.69, 1.18	0.453
Q3 (17.03, 19.88)	1.04	0.92, 1.18	0.500	1.03	0.91, 1.17	0.592	0.96	0.78, 1.18	0.679	1.11	0.84, 1.46	0.449
Q4 (19.89, 23.67)	1.01	0.90, 1.15	0.837	0.99	0.87, 1.12	0.868	0.86	0.70, 1.06	0.160	0.99	0.75, 1.32	0.954
Q5 – High (23.68, 71.18)	0.96	0.85, 1.08	0.484	0.98	0.85, 1.11	0.742	0.79	0.64, 0.98	**0.032**	1.10	0.82, 1.49	0.531
Among Children who did not change their residential postcode between 2012 -2016 (n= 2,157)												
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (13.52, 17.02)	0.96	0.82, 1.12	0.604	1.02	0.88, 1.21	0.724	0.83	0.63, 1.10	0.192	0.88	0.62, 1.24	0.460
Q3 (17.03, 19.88)	0.99	0.84, 1.16	0.872	1.02	0.87, 1.21	0.754	0.86	0.65, 1.13	0.282	0.99	0.70, 1.40	0.955
Q4 (19.89, 23.67)	0.95	0.81, 1.11	0.536	0.97	0.82, 1.15	0.717	0.89	0.68, 1.18	0.428	1.13	0.79, 1.64	0.492
Q5 – High (23.68, 71.18)	0.91	0.78, 1.07	0.297	0.96	0.80, 1.14	0.629	0.82	0.62, 1.09	0.171	1.15	0.78, 1.69	0.478

#IVdays = total number of days on IV antibiotics (home and hospital IV courses) received for a PEx during the twelve months preceding the annual review in 2016

PEx = Pulmonary exacerbation

* Adjusted for age, sex, absolute BMI, lung function (ppFEV$_1$), *P. aeruginosa*, genotype and deprivation

Table 7.4: Zero-inflated negative binomial regression analysis between NO_2 and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB

	#IVdays											
	Count model coefficient						Zero-inflation model coefficient					
	Unadjusted			Adjusted*			Unadjusted			Adjusted*		
Characteristic	**exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Ex(β)**	**95% CI**	**p-value**
All adult (≥ 16 years) (n= 5,203)												
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref			Ref			Ref			Ref		
Q2 (13.52, 17.02)	0.92	0.84, 1.01	0.071	0.95	0.87, 1.02	0.174	1.00	0.84, 1.20	0.974	1.04	0.84, 1.29	0.702
Q3 (17.03, 19.88)	0.96	0.88, 1.06	0.432	0.97	0.90, 1.06	0.511	0.99	0.83, 1.18	0.871	0.97	0.79, 1.20	0.802
Q4 (19.89, 23.67)	0.97	0.89, 1.06	0.520	0.91	0.84, 9.94	**0.035**	1.07	0.91,1.29	0.394	1.29	1.04, 1.60	**0.020**
Q5 – High (23.68, 71.18)	0.96	0.88, 1.05	0.386	0.92	0.85, 1.01	0.067	1.12	0.94, 1.33	0.217	1.24	0.99, 1.55	0.057
Among adults who did not change their residential postcode between 2012 -2016 (n= 2,891)												
NO_2 (2015 model) (µg/m^3)												
Q1 – Low (1.67,13.51)	Ref			Ref			Ref			Ref		
Q2 (13.52, 17.02)	0.94	0.83, 1.06	0.325	0.95	0.85, 1.06	0.383	1.01	0.80, 1.27	0.958	1.10	0.84, 1.45	0.489
Q3 (17.03, 19.88)	1.01	0.89, 1.14	0.848	1.00	0.90, 1.12	0.961	0.92	0.73, 1.17	0.516	0.99	0.75, 1.32	0.958
Q4 (19.89, 23.67)	1.02	0.90, 1.16	0.760	0.92	0.82, 1.03	0.158	1.05	0.83, 1.33	0.682	1.40	1.05, 1.86	**0.021**
Q5 – High (23.68, 71.18)	0.96	0.84, 1.09	0.523	0.89	0.78, 1.00	0.054	1.04	0.82, 1.32	0.737	1.34	0.98, 1.81	0.059

#IVdays = total number of days on IV antibiotics (home and hospital IV courses) received for a PEx during the twelve months preceding the annual review in 2016 PEx = Pulmonary exacerbation;

* Adjusted for age, sex, BMI percentile, lung function (ppFEV$_1$), *P. aeruginosa*, genotype and deprivation

Table 7.5: Zero-inflated negative binomial regression analysis between $PM_{2.5}$ and #IVdays in children (0-15 years) with CF in the UK CF Registry in GB

	#IVdays											
	Count model coefficient						Zero-inflation model coefficient					
	Unadjusted			Adjusted			Unadjusted			Adjusted		
Characteristic	**exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Ex(β)**	**95% CI**	**p-value**
All children (0-15 years) (n= 3,762)												
$PM_{2.5}$ (2015 model) (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (8.56, 9.41)	0.90	0.80, 1.02	0.108	0.87	0.77, 0.99	**0.029**	0.88	0.71, 1.09	0.248	0.99	0.75, 1.31	0.973
Q3 (9.42, 9.87)	0.88	0.78, 0.99	**0.044**	0.90	0.79, 1.02	0.102	0.79	0.64, 0.97	**0.028**	0.94	0.71, 1.25	0.676
Q4 (9.88, 10.27)	0.88	0.78, 1.00	0.054	0.88	0.78, 1.00	0.054	0.88	0.71, 1.09	0.248	1.04	0.78, 1.37	0.804
Q5 – High (10.28, 13.87)	0.80	0.71, 0.90	**<0.001**	0.82	0.72, 0.93	**0.002**	0.72	0.58, 0.89	**0.002**	0.92	0.70, 1.22	0.559
Among Children who did not change their residential postcode between 2012 -2016 (n= 2,157)												
$PM_{2.5}$ (2015 model) (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (8.56, 9.41)	0.81	0.69, 0.96	**0.015**	0.79	0.67, 0.93	0.005	0.97	0.73, 1.28	0.812	1.25	0.87, 1.79	0.229
Q3 (9.42, 9.87)	0.82	0.70, 0.96	**0.014**	0.86	0.72, 1.01	0.067	0.76	0.58, 1.01	0.060	0.98	0.68, 1.40	0.892
Q4 (9.88, 10.27)	0.86	0.74, 1.02	0.080	0.85	0.72, 0.99	0.049	0.97	0.73, 1.28	0.820	1.25	0.87, 1.78	0.219
Q5 – High (10.28, 13.87)	0.78	0.66, 0.91	**0.002**	0.79	0.67, 0.94	0.006	0.79	0.59, 1.05	0.106	1.05	0.73, 1.51	0.810

#IVdays = total number of days on IV antibiotics (home and hospital IV courses) received for a PEx during the twelve months preceding the annual review in 2016
PEx = Pulmonary exacerbation
* Adjusted for age, sex, BMI percentile, lung function (ppFEV$_1$), *P. aeruginosa*, genotype and deprivation

Table 7.6: Zero-inflated negative binomial regression analysis between $PM_{2.5}$ and #IVdays in adults (≥ 16 years) with CF in the UK CF Registry in GB

	#IVdays											
	Count model coefficient						Zero-inflation model coefficient					
	Unadjusted			Adjusted*			Unadjusted			Adjusted*		
Characteristic	**exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Exp(β)**	**95% CI**	**p-value**	**Ex(β)**	**95% CI**	**p-value**
All adult (≥ 16 years) (n= 5,203)												
$PM_{2.5}$ (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref			Ref	-		Ref		
Q2 (8.56, 9.41)	1.08	0.98, 1.18	0.119	1.10	1.01, 1.19	0.024	0.69	0.58, 0.82	<0.001	0.61	0.50, 0.76	<0.001
Q3 (9.42, 9.87)	1.14	1.04, 1.25	0.005	1.10	1.01, 1.19	0.029	0.79	0.66, 0.94	0.007	0.85	0.69, 1.04	0.112
Q4 (9.88, 10.27)	1.12	1.02, 1.23	0.014	1.11	1.03, 1.21	0.009	0.76	0.64, 0.90	0.002	0.76	0.62, 0.94	0.010
Q5 – High (10.28, 13.87)	1.11	1.01, 1.21	0.036	1.08	0.89, 1.17	0.087	0.95	0.80, 1.13	0.558	1.07	0.87, 1.32	0.506
Among adults who did not change their residential postcode between 2012 -2016 (n= 2,891)												
$PM_{2.5}$ (µg/m^3)												
Q1 – Low (3.20, 8.55)	Ref	-		Ref	-		Ref	-		Ref	-	
Q2 (8.56, 9.41)	1.15	1.02, 1.30	0.028	1.13	1.02, 1.26	0.025	0.69	0.55, 0.87	0.002	0.66	0.50, 0.87	0.003
Q3 (9.42, 9.87)	1.17	1.03, 1.33	0.014	1.08	0.96, 1.21	0.159	0.80	0.64, 1.00	0.054	0.96	0.72, 1.25	0.741
Q4 (9.88, 10.27)	1.18	1.04, 1.33	0.011	1.10	0.99, 1.24	0.082	0.78	0.62, 0.98	0.035	0.87	0.66, 1.14	0.303
Q5 – High (10.28, 13.87)	1.08	0.95, 1.23	0.262	1.00	0.89, 1.13	0.942	0.96	0.77, 1.22	0.758	1.14	0.87, 1.51	0.345

#IVdays = total number of days on IV antibiotics (home and hospital IV courses) received for a PEx during the twelve months preceding the annual review in 2016
PEx = Pulmonary exacerbation; $ppFEV_1$ = percent predicted with FEV_1
* Adjusted for age, sex, absolute BMI, lung function ($ppFEV_1$), *P. aeruginosa*, genotype and deprivation

7.2 Chapter III Correlation plot

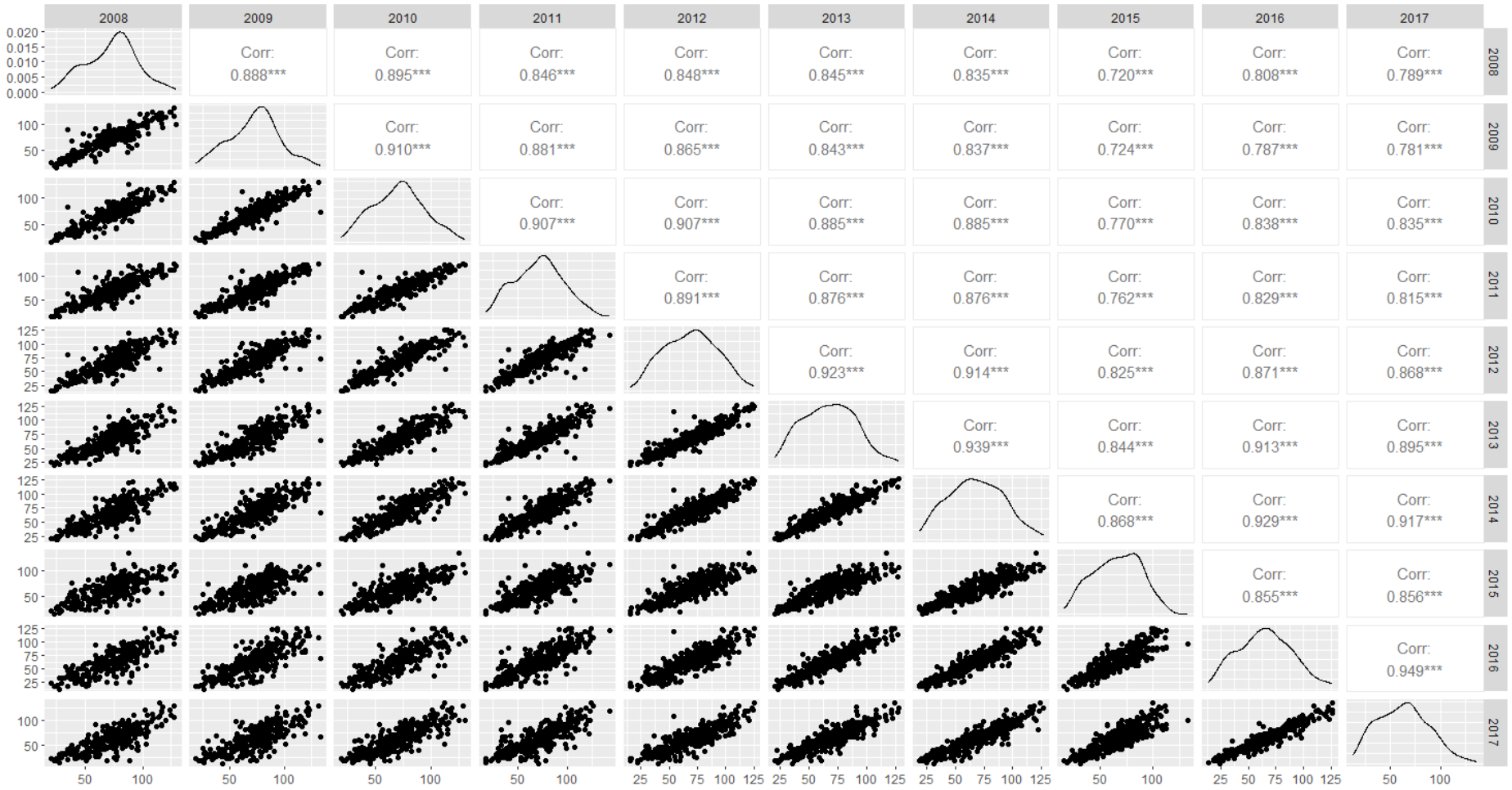


Figure 7.17: Correlation between mean ppFEV$_1$ between 2008 and 2017

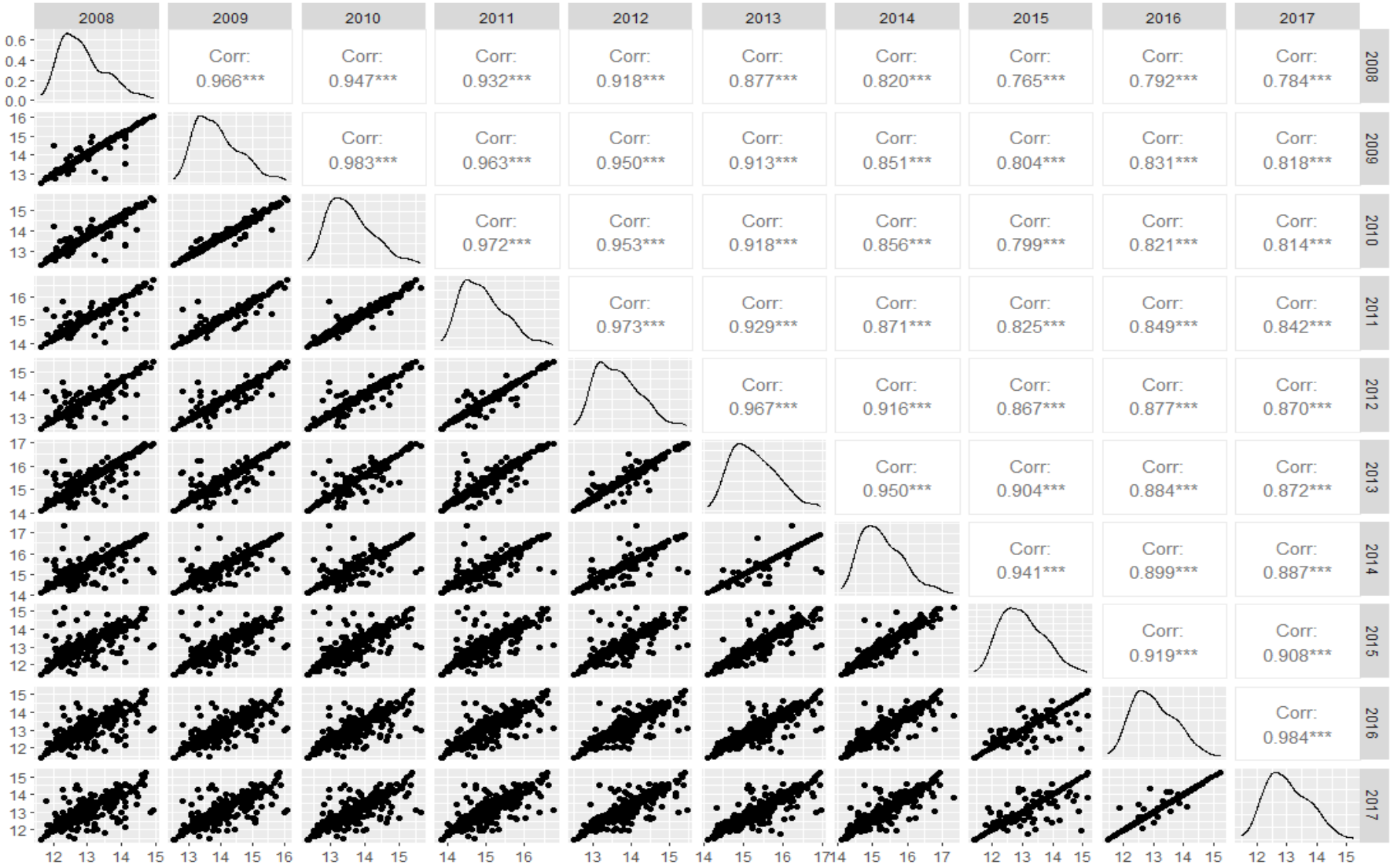


Figure 7.18: Correlation between annual $PM_{2.5}$ measurement between 2008 and 2017

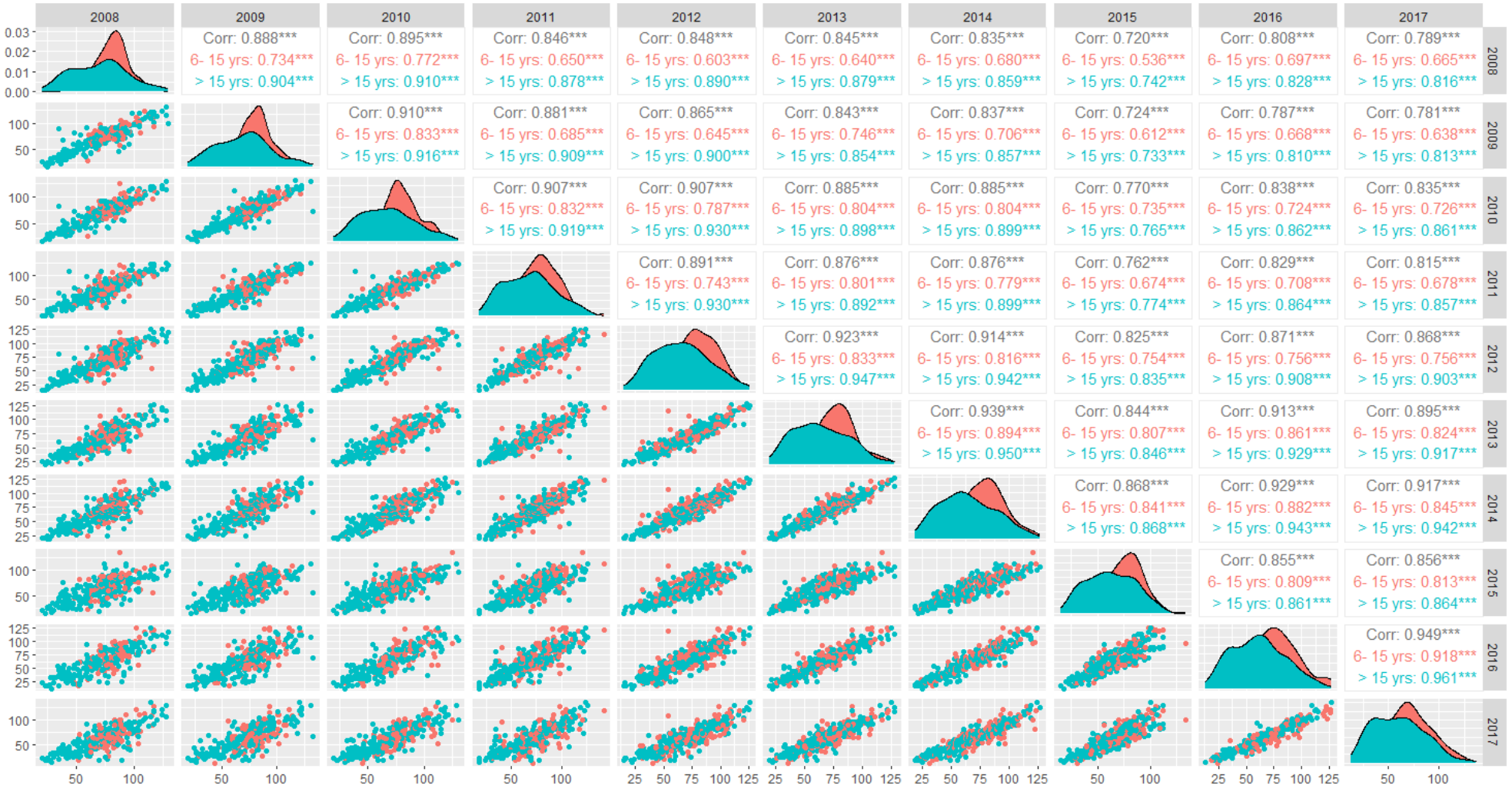


Figure 7.19: Comparison of correlation between ppFEV$_1$ between 2008 and 2017 by age group