UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Clinical and Experimental Science

Volume 1 of 1

Delivering health care for patients with primary ciliary dyskinesia: diagnosis and life-long care

by

Bruna Rubbo

Thesis for degree of Doctor of Philosophy

June 2020

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE

Infection, Inflammation and Immunity; Clinical and Experimental Science School

Thesis for degree of Doctor of Philosophy

DELIVERING HEALTH CARE FOR PATIENTS WITH PRIMARY CILIARY DYSKINESIA: DIAGNOSIS AND LIFE-LONG CARE

Bruna Rubbo

Primary ciliary dyskinesia (PCD) is a rare genetic disease that affects the function of cilia lining the upper and lower respiratory airways. High-speed video microscopy analysis (HSVA) is the only clinically available diagnostic test to assess ciliary function; however, European and North American diagnostic guidelines could not recommend its use as a confirmatory test due to insufficient evidence. The inability of cilia to beat in a coordinated manner can lead to varying degrees of progressive lung disease, recurrent otitis media with hearing loss and chronic rhinosinusitis. Disease expression is likely linked to genotype, with some genes responsible for a more severe phenotype. There is a lack of disease-specific validated clinical outcome measures to monitor disease progression and for use in clinical trials and prospective cohorts. Furthermore, service delivery models vary from country to country, yet there have been no comparisons on the effect of these differences in delivery of care for patients with PCD.

Objectives: This work sets out to examine different aspects of diagnosis and management of PCD patients by determining: **a)** the accuracy and reliability of HSVA to diagnose PCD, **b)** genotype-phenotype associations, **c)** appropriate clinical outcome measures to monitor disease progression and for use in research, and **d)** differences in service delivery that could affect disease outcomes.

Methods: We calculated the sensitivity, specificity, inter- and intra-observer reliability of HSVA to diagnose PCD compared to both European diagnostic guidelines and the multidisciplinary team clinical decisions **(a)**. In order to investigate genotype-phenotype associations, we collected diagnostic, clinical and genetic data from three countries. We then applied a novel data-driven clustering method (i.e. topological data analysis) to generate models that would highlight underlying patterns in the phenotypic data and guide the selection of two genes for standard hypothesis testing **(b)**. We conducted a scoping review in order to identify clinical outcome

measures that have been used in PCD research and to recommend the use of a set of outcomes in future longitudinal studies **(c)**. Finally, we performed interviews with PCD experts from various countries to develop a survey that would quantify similarities and differences between service delivery models. These were compared to results from a survey conducted ten years ago **(d)**.

Results: We found that HSVA was both sensitive and specific to diagnose PCD, with high inter- and intra-observer agreement when performed by experienced scientists **(a)**. This was the first powered and blinded study to assess HSVA in isolation compared to the current evidence-based European diagnostic guidelines.

Our topological models revealed a cluster of patients with mutations on the *CCDC39* gene that had more severe respiratory phenotype, with statistically significantly lower FEV_1 z-scores at diagnosis and higher proportion of history of neonatal respiratory distress **(b)**. We also found a cluster of patients with *DNAH11* mutations that had statistically significant milder respiratory phenotype. This was the first large scale multi-centred European study to investigate phenotype-genotype associations in PCD across multiple diagnostic and clinical parameters, using a novel data-driven approach.

The scoping review identified a variety of clinical outcome measures that have been used in PCD research, with significant differences in measurement and reporting of outcomes between studies **(c)**. It also highlighted the lack of standardisation in selection, definition and application of outcomes across studies. We recommended using disease-specific validated outcome measures such as QOL-PCD in future randomised controlled trials in order to generate comparable results.

Comparisons between service delivery models revealed that diagnostic testing for PCD has become more standardised and centralised compared to ten years ago **(d)**. However, management still varies considerably depending on geographical region and size of the PCD centre, with larger centres prioritising therapies that have been shown to be effective in PCD such as airway clearance and nasal rinsing.

Discussion: The findings from these studies provide evidence to improve the delivery of health care for patients with PCD. We showed that HSVA can play an important part in the diagnostic pathway. Our genotype-phenotype association study provided evidence for early and more aggressive intervention in patients with specific genetic mutations. The adoption of a core set of clinical outcome measures will generate the evidence needed for the development of management guidelines, the importance of which were highlighted when comparing service delivery models between countries.

Table of Contents

Table of Contents ..... **i**

Table of Tables ..... **vii**

Table of Figures ..... **ix**

Academic Thesis: Declaration Of Authorship **11**

Definitions and Abbreviations **13**

Acknowledgments **15**

Chapter 1 Introduction **17**

1.1 Thesis outline 17

1.2 Author's contribution 18

Chapter 2 Literature review **21**

2.1 Rare diseases 21

2.1.1 Challenges for rare diseases 21

2.1.2 The European strategy: a short history 22

2.1.3 Specialised Centres of Reference 24

2.2 Primary ciliary dyskinesia 25

2.2.1 Cilia ultrastructure and function 25

2.2.2 PCD genetics 26

2.2.3 Clinical manifestations 27

2.3 Referral for diagnostic testing 29

2.4 PCD diagnostic testing 31

2.4.1 Nasal nitric oxide 31

2.4.2 High-speed video analysis 33

2.4.3 Transmission electron microscopy 34

2.4.4 Genetic testing 35

2.4.5 Other diagnostic tests 35

2.4.6 Diagnostic guidelines 36

2.4.7 Standardised reporting for diagnostic testing 38

2.5 PCD management 39

2.5.1 Lower airway management 40
2.5.2 Upper airway management 44
2.5.3 Cardiac management 45
2.5.4 Fertility 46
2.5.5 Potential future therapies 46
2.6 Rationale and aims 49
Chapter 3 General methods 51
3.1 Data protection and research ethics 51
3.2 Research setting 52
3.2.1 PCD diagnostic service in England 52
3.2.2 PCD management service in England 54
3.3 Data collection, management and analyses 55
Chapter 4 Accuracy of high-speed video analysis to diagnose PCD 57
4.1 Introduction 57
4.1.1 High-speed video analysis as a diagnostic tool 57
4.1.2 Aims 58
4.2 Methodology 58
4.2.1 Patient population and diagnostic decisions in the clinical setting 58
4.2.2 Inclusion and exclusion criteria for selection of study samples 59
4.2.3 Study population and data 59
4.2.4 Selection of reference standards 60
4.2.5 Statistical analyses 63
4.2.6 Pilot study 65
4.3 Results 65
4.3.1 Accuracy of HSVA compared to the ERS defined outcomes (i) 68
4.3.2 Accuracy of HSVA compared to MDT decision (ii) 70
4.3.3 Intra-observer reliability for each scientist 72
4.3.4 Inter-observer reliability for each diagnostic outcome 72
4.3.5 Evaluation of PCD cases with a genetic diagnosis 72
4.4 Discussion 73

4.4.1 Accuracy of HSVA to diagnose PCD 73
4.4.2 Reliability of HSVA to diagnose PCD 74
4.4.3 Implications to diagnostics and clinical practice 75
4.4.4 Strengths and limitations 75
4.5 Conclusions 76

Chapter 5 Genotype-phenotype associations in PCD using topological data analysis . 79

5.1 Introduction 79
5.2 Study aims 81
5.3 Methodology 81
5.3.1 Data collection 81
5.3.2 Inclusion criteria and study design 82
5.3.3 Data classification tree 82
5.3.4 Topological data analysis 83
5.3.5 Selection of topological model 85
5.3.6 Statistical analysis 86
5.4 Results 87
5.4.1 Discovery dataset 87
5.4.2 Validation dataset 97
5.4.3 Hypothesis testing of patients with *CCDC39* and *DNAH11* mutations 104
5.5 Discussion 104
5.5.1 Main findings and comparison with previous literature 105
5.5.2 Potential implications to clinical practice 106
5.5.3 Strengths and limitations 106
5.6 Conclusions 107

Chapter 6 Clinical outcome measures for use in PCD research, a scoping review 109

6.1 Introduction 109
6.1.1 Disease-specific outcome measures 110
6.1.2 Scoping reviews 110
6.1.3 Study aims 111
6.2 Methodology 111

6.2.1 Search strategy 111
6.2.2 Inclusion and exclusion criteria 111
6.2.3 Manuscript selection 112
6.2.4 Definition of outcome measures and classification into subgroups 113
6.2.5 Statistical analysis 113
6.3 Results 116
6.3.1 Study characteristics 117
6.3.2 Outcome measures reported 118
6.3.3 Standardised definitions 119
6.3.4 Lung function 120
6.3.5 Imaging 122
6.3.6 Microbiology 124
6.3.7 Other outcome measures 125
6.3.8 Comparison between outcome measures 126
6.3.9 Randomised controlled trials 127
6.4 Discussion 129
6.4.1 Limitations 131
6.4.2 Recommendations 132
6.5 Conclusions 135
Chapter 7 Comparison of service delivery models for PCD 137
7.1 Introduction 137
7.1.1 Literature review 137
7.1.2 Gap in the literature 138
7.1.3 Study aims 139
7.2 Methodology 139
7.2.1 Mixed methods and theoretical approach 139
7.2.2 Choice of data collection method 140
7.2.3 Study design and documents 141
7.2.4 Qualitative data analysis 141
7.2.5 Survey development and structure 142

7.2.6 Study population and quantitative data collection 142
7.2.7 Choice of typology for country classification 143
7.2.8 Quantitative data analysis 143
7.3 Results 146
7.3.1 Qualitative research 146
7.3.2 Survey demographics 146
7.3.3 Complete survey analysis, stratified by geographical regions 149
7.3.4 Complete survey analysis by size of PCD centre 153
7.3.5 Comparison with previous survey 157
7.4 Discussion 161
7.4.1 Diagnostic strategies 161
7.4.2 Management strategies 161
7.4.3 Referrals and funding structure 163
7.4.4 Potential implications 163
7.4.5 Strengths and limitations 164
7.5 Conclusions 165

Chapter 8 General discussion 167

8.1 Diagnostic phenotype 168
8.2 Diagnostic strategy 169
8.3 Disease phenotype 170
8.4 Management strategy 171
8.5 Strategy for delivery of care 172
8.6 Strengths and limitations 173
8.6.1 Strengths 173
8.6.2 Limitations 174
8.7 Future directions for PCD clinical services 176
8.8 Future research directions 176
8.9 General conclusions 178

Appendix A Accuracy of high-speed video analysis to diagnose PCD 181

A.1 Variable coding 181

A.2 Variable definition 185
A.3 Additional results 190

Appendix B Previously described genotype-phenotype associations in PCD..... 195

Appendix C Clinical outcome measures for use in PCD: studies included in the scoping review 203

C.1 Study characteristics 203
C.2 Outcome measures definitions 239
C.3 Computed tomography scoring scales 302

Appendix D Comparison of service delivery models for PCD: study pack documents ... 305

D.1 Participant information sheet 305
D.2 Participant consent form 308

Appendix E Comparison of service delivery models for PCD: interview guide and complete survey 309

E.1 Interview guide 309
E.2 Service delivery complete survey 311
E.3 Typology for service delivery studies 327

List of References 337

Table of Tables

Table 1. Summary of current knowledge, recent advances and remaining barriers in PCD 47

Table 2. Equipment used at the PCD centres for high-speed video analysis, nasal nitric oxide and electron microscopy. 53

Table 3 Clinical characteristics of study participants.....66

Table 4 Aggregated diagnostic study outcomes by the three scientists compared to the diagnostic outcome defined by ERS guidelines 69

Table 5. Aggregated diagnostic study outcomes by the three scientists compared to the original diagnostic decision made by the MDT.....71

Table 6. Diagnostic characteristics of patients in the discovery group, stratified by predefined gene groups. 88

Table 7. Clinical characteristics of patients in the discovery group, stratified by predefined gene groups. 91

Table 8. Diagnostic characteristics of patients in the validation group, stratified by predefined gene groups. 97

Table 9. Clinical characteristics of patients in the validation group, stratified by predefined gene groups. 100

Table 10. Summary of study characteristics of crossover randomised controlled trials included in this scoping review. 128

Table 11. Summary of recommendations on clinical outcome measures for use in PCD research. 134

Table 12. Demographic characteristics of survey participants, stratified by size of PCD diagnostic centre.....148

Table 13. Average number (%) of patients referred, diagnosed and managed at PCD centres per year, stratified by region 156

Table 14. Distribution of centres that completed each survey, stratified by UN European region155

Table of Tables

Table 15. Data dictionary for clinical and diagnostic data included in the study depicted in Chapter 4. 181

Table 16. Data dictionary for outcome variables included in the study depicted in Chapter 4 185

Table 17. Diagnostic outcomes decisions made by each of the three high-speed video analysis (HSVA) scientists (*i.e.* scientist A-C, *n*=120 for each) compared to the ERS guidelines (25). 190

Table 18. Diagnostic outcomes decisions made by each of the three high-speed video analysis (HSVA) scientists (*i.e.* scientist A-C, *n*=120 for each) compared to the original diagnostic decision made by the MDT 191

Table 19. Diagnostic outcomes decisions from a random subset of patient-samples (*n*=20) by each of the three scientists (*i.e.* scientist A-C) using high-speed video analysis (HSVA) at two different time points (*i.e.* original study analysis and re-analysis after one year)..... 193

Table 20. Summary of the phenotype-genotype relationships described in the PCD literature.195

Table 21. Characteristics of studies included in the clinical outcome measures scoping review203

Table 22. Definition of outcome measures included in the clinical outcome measure scoping review, stratified by study outcome and population descriptor. 239

Table 23 Comparison between computed tomography (CT) scoring scales used in PCD studies included in the clinical outcome measure scoping review. 302

Table 24. Summary of typologies for country classification..... 327

Table of Figures

Figure 1. Schematic depicting normal ultrastructure of a cross-section of motile cilium. 26

Figure 2. PCD classification according to the ERS guidelines 37

Figure 3. Schematic depicting the use of the ERS guidelines as reference standard to assess the accuracy of high-speed video analysis (HSVA) 61

Figure 4. Schematic depicting the use of clinical multidisciplinary team outcome as reference standard to assess the accuracy of high-speed video analysis (HSVA) 62

Figure 5. Schematic depicting the diagnostic algorithms used to determine 'positive' and 'negative' outcomes used to calculate sensitivity and specificity 64

Figure 6. Schematic of stratification of PCD-causative genes into groups according to the ciliary components they encode 82

Figure 7. The three fundamental principles that govern topology 84

Figure 8. Schematic illustrating the generation of topological networks in Ayasdi 85

Figure 9. Distribution of genes mutated in the discovery dataset 87

Figure 10. Topological discovery model. Models A-C are coloured by the following features: A. gene group, B. transmission electron microscopy (TEM) results, and C. ciliary beat pattern (CBP) by high-speed video analysis (HSVA) 94

Figure 11. Topological discovery model. Models A-F are coloured by the following features: A. gene group, B. FEV_1 z-scores, C. history of neonatal respiratory distress syndrome (NRDS), D. *CCDC39* mutations, E. *DNAH11* mutations, F. *DNAH5* mutations 95

Figure 12. Topological discovery model. Models A-D are coloured by the following features: A. gene group, B. history of rhinitis, C. history of glue ear, D. cardiac situs 96

Figure 13. Distribution of genes mutated in the validation dataset 97

Figure 14. Topological validation model. Models A-F are coloured by the following features: A. Gene group, B. FEV_1 z-scores, C. history of neonatal respiratory distress syndrome (NRDS), D. *CCDC39* mutations, E. *DNAH11* mutations, F. *DNAH5* mutations 103

Table of Figures

Figure 15. PRISMA flow diagram for the selection of studies reviewed and included in the systematic review. 117

Figure 16. Number of PCD patients per study by year of publication, stratified by country 118

Figure 17. Number of studies that reported outcome measures in PCD as either study outcome or population descriptor 119

Figure 18. Number of studies that reported on each pathogen, stratified by age group (<18 and ≥18 years of age, or not differentiated). 120

Figure 19. Studies that presented FEV_1 as study outcome or population descriptor 121

Figure 20. High-resolution computed tomography (HRCT) and CT outcome measures from 20 studies that reported on HRCT/CT scans as study outcome. 124

Figure 21. Correlations or associations between outcomes measures. 127

Figure 22. Average number of patients referred, diagnosed and managed at PCD centres per year, grouped by country and region. 145

Figure 23. Mapping of themes derived from the interviews to questions added to the survey. 147

Figure 24. Map highlighting the countries included in the survey. 150

Figure 25. Diagnostic (top) and management (bottom) techniques stratified by geographical regions, described by participants of the study survey (2019). 152

Figure 26. Use of nasal rinsing, hypertonic saline and antibiotic prophylaxis across management centres, stratified by size of management centre. 154

Figure 27. Comparison of diagnostic tests available according to the 2009 and 2019 surveys, stratified by European region. 159

Figure 28. Comparison of management strategies according to the 2009 and 2019 surveys, stratified by European region. 160

Figure 29. Venn diagram summarising the service delivery framework adopted in this thesis. 167

Academic Thesis: Declaration Of Authorship

I, Bruna Rubbo, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Delivering health care for patients with primary ciliary dyskinesia: diagnosis and life-long care

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Rubbo B and Lucas JS. Clinical care for primary ciliary dyskinesia: current challenges and future directions. *Eur Respir Rev* Sep 2017, 26 (145) 170023; DOI: 10.1183/16000617.0023-2017

Rubbo B, Shoemark A, Jackson CL, Hirst R, Thompson J, Hayes J, Frost E, Copeland F, Hogg C, O'Callaghan C, Reading I, Lucas JS. Accuracy of high-speed video analysis to diagnose primary ciliary dyskinesia. *Chest* 2019;155(5):1008-1017

Signed: .....

Date: .....

Definitions and Abbreviations

ATS: American Thoracic Society

BMI: body mass index

CBF: ciliary beat frequency

CBP: ciliary beat pattern

CF: cystic fibrosis

CHD: congenital heart defect

CI: confidence interval

COPD: chronic obstructive pulmonary disease

(N-)DRC: (nexin-)dynein regulatory complex

ENT: ear, nose, throat specialist

ERN: European Reference Network

ERS: European Respiratory Society

EU: European Union

FEV_1: forced expiratory volume in the first second

FVC: forced vital capacity

HRCT: high-resolution computed tomography

HRQoL: health-related quality of life

HSVA: high-speed video analysis

IDA: inner dynein arm

IQR: inter-quartile range

LCI: lung clearance index

LRI: Leicester Royal Infirmary

Definitions and Abbreviations

MBW: multiple breath washouts

MDT: multidisciplinary team

MTD: microtubular disarrangement

MRI: magnetic resonance image

NHS: National Health Service

nNO: nasal nitric oxide

NDRS: neonatal respiratory distress syndrome

ODA: outer dynein arm

OM: otitis media

PCD: primary ciliary dyskinesia

RBH: Royal Brompton Hospital

RCT: randomised controlled trial

rhDNase: Recombinant human deoxyribonuclease I

SD: standard deviation

SGRQ: St George's Respiratory Questionnaire

SOP: standard operating procedure

TDA: topological data analysis

TEM: transmission electron microscopy

UHS: University Hospital Southampton

Acknowledgments

This thesis would not have been completed without the support of various people. In the academic sphere, my three supervisors were indispensable.

My primary supervisor, Prof Jane Lucas, was an invaluable source of knowledge and inspiration. Her dedication to advancing research in the field of PCD, and her constant encouragement were fundamental to everything I have achieved in the last five years. I could not have asked for a better mentor, supervisor and line manager. Dr Camille Parsons provided essential support for the statistical analyses carried out throughout the thesis. Dr Laura Behan was a constant source of assistance, mentorship and friendship throughout the five years, a true role-model.

All the studies included in this thesis were a direct result of collaboration with colleagues from different specialities and countries. I was incredibly fortunate to work with the fantastic PCD team at Southampton General Hospital during these years. Their dedication to research and clinical care is truly inspiring. The BEAT-PCD network was crucial to the completion of all the projects included in this thesis - in particular, Dr Kuehni and her team in Switzerland. I am truly blessed to have been part of such a collaborative and productive network.

In the personal sphere, my family and friends were priceless in offering support and encouragement throughout the last five years. This thesis is dedicated to my close family. My sister was a constant companion, on holidays, football matches, music concerts, film festivals, and social activities in general. She certainly kept me centred and sane during all these years, particularly through the hard times. I would not have been able to complete this thesis without her unending support – and her expertise in graphic design. Equally, my mother and father were fundamental to the completion of this thesis. Their love, support and belief in me were unwavering, not only in the last five years but throughout my entire life.

I have been blessed with many supportive friends, in fact too many to name here. They all know who they are and how important they are to me. I would like to thank, in particular, Dr Bram Duyx, for being such a great friend and for helping me devise a concrete plan to reach the finish line; Dr Heena Mistry, for making my life in Southampton so enjoyable; Amanda Harris and Dr Corine Driessens, for hours of insightful conversations; Billy Hillier, for always making me laugh; Johnny Goodwin, and Laura Behan's family, for their hospitality; and Dr Eva Zoubek, Joshua Wallace, Iris Kouneni, and Alexis Dallas for so many great conversations and memories.

In loving memory of my cats, Billy and Joe.

Chapter 1 **Introduction**

1.1 Thesis outline

This thesis is comprised of four related projects that cover different aspects of service delivery for primary ciliary dyskinesia (PCD), a rare genetic disease. Chapter 2 provides a comprehensive literature review, focusing on the complexities of delivery of care for rare diseases, and the pathophysiology, diagnosis and management of PCD. It highlights the gaps in current knowledge and concludes with the rationale and aims of the research studies included in this thesis.

Chapter 3 details the methods and setting used to collect clinical and diagnostic data. In Chapter 4, the accuracy and reliability of high-speed video analysis (HSVA), a commonly used diagnostic test, was determined through a multicentre retrospective study. Findings provide evidence that can be used to update current diagnostic guidelines and better inform the positioning of this test in the diagnostic pathway.

In Chapter 5, we investigate clusters of phenotypic characteristics in PCD patients and their association with different gene mutations using a novel method, topological data analysis coupled with machine learning algorithms, to select candidate genes for traditional hypothesis testing. I describe the phenotype of mutations in two genes, *CCDC39* and *DNAH11*. These findings can inform the monitoring and treatment of patients with these mutations in the future.

Chapter 6 describes results from a wide-range scoping review on clinical outcome measures used in PCD research. It focuses on the heterogeneity of definitions, measurements and reporting of outcome measures used in the current literature. Findings from this chapter can assist in the development of expert-led consensus on a core set of outcome measures that should be used in future longitudinal studies such as prospective cohorts and randomised controlled trials (RCTs).

Chapter 7 explores differences in service delivery models for PCD in different countries using mixed-methods research. I conducted interviews with PCD specialists to develop a survey that was distributed to invited experts at PCD diagnostic and management centres worldwide. Results from the survey were compared to a similar cross-sectional study conducted a decade ago to assess trends on diagnostic and management techniques across Europe. Results presented in this chapter highlight shifts from decentralised to centralised models of care for PCD and the impact of evidence-based diagnostic guidelines.

Finally, Chapter 8 and Chapter 9 discuss the overarching findings from the projects included in this thesis, focusing on the following themes: diagnostic phenotype, diagnostic strategy, disease

phenotype, management strategy, and strategy for delivery of care. I also highlight common strengths and limitations and provide recommendations for future research studies in delivery of care for PCD.

1.2 Author's contribution

I was the lead author for the manuscripts based on Chapter 4, Chapter 5 (joint first author with Dr Amelia Shoemark) and Chapter 6 (joint first author with Dr Florian Gahleitner).

Findings from Chapter 4 were published in Chest on May 2019 (Volume 155, Issue 5, Pages 1008–1017, DOI: https://doi.org/10.1016/j.chest.2019.01.036). I assisted Prof Jane Lucas in preparing the grant proposal by actively contributing to the study design and data analyses plan. I liaised with collaborators in London and Leicester and developed the timetable for milestones and key deliverables. With the guidance of Prof Lucas, I submitted the application to NIHR and dealt with any issues raised by the reviewing panel. I coordinated the project through constant liaison with technicians and scientists from the three PCD diagnostic services in England. I drafted the study protocol and SOP for video selection (jointly with James Thompson). I conducted all statistical analyses, under the supervision of statistician Dr Isabel Reading, and ensured the study was completed in a timely manner. I presented the findings from this study at the 2017 European Respiratory Society Congress (Milan, Italy) in a poster discussion session and as an oral presentation at the 4th BEAT-PCD Conference (Poznan, Poland).

The study depicted in Chapter 5 has been submitted to the European Respiratory Journal and is currently under review. I assisted Dr Shoemark, Prof Lucas, Prof Claire Hogg and Dr Hannah Mitchison in developing the initial concept, study design and data analysis plan. I collected data from University Hospital Southampton (UHS) and assisted Dr Shoemark in collecting data from the Royal Brompton Hospital (RBH) in London. I coded, cleaned and re-coded data, with the exception of genetic data, which was cleaned and coded by Dr Mitchison and Dr Marie Legendre. Dr Joost Brandsma introduced me to Ayasdi, the online platform used to conduct topological data analysis. I generated and selected the topological models, which was reviewed by Prof Gunner Carlson (topological data analysis expert and co-founder of Ayasdi), and conducted hypothesis testing of selected genes, which was reviewed by Dr Camille Parsons (statistician). Dr Shoemark and I jointly drafted the manuscript.

Under the guidance of Prof Lucas and Prof Philipp Latzin, I developed the study protocol and data extraction sheet for the scoping review described in Chapter 6. I led the data extraction and was assisted by Dr Gahleitner, who collected data for half of the studies deemed eligible for inclusion. Dr Gahleitner organised data from studies describing high-resolution computed tomography

(HRCT), body plethysmography and multiple breath washouts (MBW). Jana Hueppe summarised the microbiology data. I organised data from all other outcomes, performed data analysis and developed all figures and tables included in Chapter 6.

I led and conducted all aspects of the study detailed in Chapter 7. I was guided by Dr Laura Behan on mixed-method research and qualitative analysis, including the conduct of interviews, use of appropriate software and development of the survey. Prof Lucas, Dr Behan, Prof Claudia Kuehni and Dr Parsons offered important advice on the survey development, structure and analysis of results. Hannah Wilkins and Amanda Harris piloted the survey and provided feedback. Lynn Reeves and James Thompson assisted in distributing the survey to PCD experts during the 4th BEAT-PCD Conference.

Additionally, I was the lead or co-author for the following manuscripts, which contributed to the literature review detailed in Chapter 2, the discussions in Chapter 4 and the overall discussion in Chapter 8:

Rubbo B, Best S, Hirst RA, Shoemark A, Goggin P, Carr SB et al. Clinical features and management of children with primary ciliary dyskinesia in England. *Arch Dis Child.* **2020**; archdischild-2019-317687. doi: 10.1136/archdischild-2019-317687

Shoemark A, Rubbo B, Haarman E, Hirst RA, Hogg C, Jackson CL, Nielsen KG, Papon J-F, Robinson P, Walker WT, Lucas JS. The Controversies and Difficulties of Diagnosing Primary Ciliary Dyskinesia. *Am J Respir Crit Care Med.* **2020**; 201(1):120-122

Goutaki, M., Papon, J-F., Boon, M., Carmen, C., Eber, E., Escudier, E., Halbeisen, F. S., Harris, A., Hogg, C., Honore, I., Jung, A., Karadag, B., Koerner-Rettberg, C., Legendre, M., Maitre, B., Nielsen, K., Rubbo, B., et al. Standardised clinical data from patients with primary ciliary dyskinesia (PCD): FOLLOW-PCD. Eur Respir J Open. 2020; 6: 00237-2019

Rubbo B, Shoemark A, Jackson CL, Hirst R, Thompson J, Hayes J, Frost E, Copeland F, Hogg C, O'Callaghan C, Reading I, Lucas JS. Accuracy of high-speed video analysis to diagnose primary ciliary dyskinesia. *Chest* 2019;155(5):1008-1017

Rubbo B, Goutaki M, Lucas JS. Primary ciliary dyskinesia. *Eur Respir Soc Handbook Paediat Respir Med.* **2019**

Nenna R, Hunt KA, Dassios T, Collins JJP, Rottier RJ, Liu NM, Rottier B, Goutaki M, Karadag B, Prayle A, Fernandes RM, Parisi GP, Barben J, Rubbo B et al. Key paediatric messages from the 2018 European Respiratory Society International Congress. *ERJ Open Res.* **2019**; 5(2):00241-2018

Kuehni C, Goutaki M, Rubbo B, Lucas JS. Management of primary ciliary dyskinesia: current practice and future perspectives. *Eur Respir Soc Monograph.* **2018**

Queiroz APL, Athanazio RA, Olm MAK, Rubbo B et al. Translation of the quality-of-life measure for adults with primary ciliary dyskinesia and its application in patients in Brazil. J Bras Pneumol. 2018; 45(2):e20170358

Farley H, Rubbo B et al. Proceedings of the 3rd BEAT-PCD Conference and 4th PCD Training School. BMC Proc December 2018; 12(Suppl 16):64

Best S, Shoemark A, Rubbo B et al. Risk factors for situs defects and congenital heart disease in primary ciliary dyskinesia. Thorax 2018; doi: 10.1136/thoraxjnl-2018-212104

Halbeisen F, Hogg C, Alanin MC, Bukowy-Bieryllo Z, Dase F, Duncan J, Friend A, Goutaki M, Jackson C, Keenan V, Harris A, Hirst RA, Latzin P, Marsh G, Nielsen K, Norris D, Pellicer D, Reula A, Rubbo B et al. Proceedings of the 2nd BEAT-PCD conference and 3rd PCD training school. BMC Proc 2018; 12(Suppl 2):1

Rubbo B and Lucas JS. Clinical care for primary ciliary dyskinesia: current challenged and future direction. Eur Respir Rev 2017; 26:170023

Shoemark A, Moya E, Hirst RA, Patel MP, Robson EA, Haywayrd J, Scully J, Fassad M, Lam W, Schmidts M, Dixon M, Patel-King RS, Rogers AV, Rutman A, Jackson C, Goggin P, Rubbo B et al. A high prevalence CCDC103 p.His154Pro mutation causing primary ciliary dyskinesia disrupts protein oligomerisation and is associated with normal diagnostic investigations. Thorax 2017; doi: 10.1136/thoraxjnl-2017-209999

Behan L, Rubbo B, Lucas JS, Dunn Galvin A. The patient's experience of primary ciliary dyskinesia: a systematic review. Qual Life Res 2017; doi: 10.1007/s11136-017-1564-y

Rubbo B, Behan L et al. Proceedings of the COST action BM1407 inaugural conference BEAT-PCD: translational research in primary ciliary dyskinesia - bench, bedside, and population perspectives. BMC Procs 2016; 10(S9):5-19

Behan L, Galvin AD, Rubbo B, et al. Diagnosing primary ciliary dyskinesia: An international patient perspective. Eur Respir J 2016; 48: 1096–1107

Chapter 2 **Literature review**

2.1 Rare diseases

Rare diseases are defined by the European Union (EU) as those that affect not more than 5:10 000 persons (1). However, this group of highly heterogeneous diseases is estimated to affect an average of 246 000 patients per disease in the EU (2), and these numbers are expected to increase as better diagnostic methods, coupled with improved screening and reporting, are developed and implemented.

Primary ciliary dyskinesia (PCD) is a rare disease with an estimated prevalence of 1:10 000 (3-5); however, true prevalence of the disease is still unknown largely due to underdiagnoses. The childhood population in England is estimated to be around 11.8 million. With an expected prevalence of 1 in 10 000, there should be approximately 1 180 children diagnosed with PCD; however, there were approximately 330 children diagnosed with PCD under the English National Children's PCD Management Service in 2015 (6). Higher prevalence has been reported in close-nit communities with a strong component of geographical or cultural isolation and high levels of consanguinity. For example, the incidence of PCD was 1:2 265 in the Asian British population in Northern England, of which the majority were Pakistani (7), and even higher (1:400) in subsection of the Dutch population residing in a culturally and geographically isolated fishing village in Northern Holland (8).

2.1.1 Challenges for rare diseases

Difficulties in identifying cases often result in varied or unknown prevalence across different countries. These include the lack of centralised international disease registries, limited awareness of disease by clinicians and patients, difficulties in reaching and accessing diagnostic facilities, and the need for specialist equipment for diagnostic testing.

A study comparing the prevalence of patients diagnosed with PCD per European country reported wide variations, with the highest prevalence in Cyprus (1:10 000) (4). Many rare diseases have non-specific or typical symptoms or co-exist with diseases that are more common, posing a challenge for non-specialist clinicians. This results in an increase on the diagnostic delay interval, defined as the time between the initial symptoms attributed to the disease and the diagnosis. Patients suffering from rare diseases often see multiple clinicians before being referred to a specialist (9). An international survey on the patients' perspectives on PCD diagnosis found that 35% of patients with PCD-like symptoms had over 40 clinical appointments before being referred

to a specialist for diagnostic testing, and a German survey reported approximately 30% had seen a clinician over 50 times before they were diagnosed with PCD (10, 11).

In the field of rare diseases, where there are limited number of cases in a wide geographical area, various sources of information must be linked in order to obtain sufficient data on which to base the development and selection of these indicators (i.e. administrative databases, national and regional health statistics, patient support group databases, national and regional surveys, disease registries). For many rare diseases, including PCD, there is no International Classification of Diseases (ICD) code in its current version 11, hampering direct identification of cases from electronic hospital records, hospital admission databases, discharge records, death certificates, and official statistics in general.

Disease registries can pool cases together in a standardised dataset that can be used for international comparisons. The Building Consensus and Synergies for the EU Registration of Rare Disease Patients (EPIRARE) was tasked with standardising datasets, governance and scope of an EU disease registry platform. The process of data harmonisation is also the main objective of the RD-Connect project, an integrated platform connecting registries, biobanks and clinical bioinformatics for rare diseases. Additionally, EUROPLAN seeks to develop a common set of indicators to standardise the use of data among the 28 EU Member States, increasing data shareability and comparability.

For many rare diseases, there are no disease-specific therapies available and treatment is often symptom-based. The reduced number of novel disease-specific medications is partly due to the limited potential market-size, as the drug development process requires large investments in time, money and resources from the pharmaceutical sector. Conditions that only affect a reduced number of patients offer small financial returns compared to diseases that are more common. Potentially more promising strategies include investigating new uses for already approved and commercialised medicines.

2.1.2 The European strategy: a short history

The challenges outlined in 2.1.1 emphasise the need for a common approach for rare diseases by setting out strategies at the highest possible level. The first European legislation on rare diseases was the 1999 Regulation (European Commission) 141/2000 on Orphan Medicinal Products (12). This important piece of legislation standardised the definition of rare diseases at the European level and formed the basis upon which further legislation was built.

In 2004, a Working Group on European Centres of Reference was established with the aim of providing expert advice to the European Commission's Public Health Directorate on the technical and scientific aspects of establishing centres of reference. The final report, published in 2005, highlighted the lack of a common definition of Centre of Reference on Rare Diseases between EU Member States, as well as differences in the rigour each country applied to the current definition of rare diseases. They also found that the number of European Centres of Reference was associated with the organisational structure of healthcare systems (i.e. centralised versus decentralised approach), as opposed to population size. The Working Group provided recommendations on criteria for selecting new centres and identified six areas for future work: a) mapping existing expert centres, b) networking between centres, c) developing and managing shared case management systems through telemedicine, d) online diagnosis, meetings and shared repository of cases, e) development of a formal process for designation of European Centre of Reference, and f) dissemination of information to stakeholders (2).

The European Commission developed the EU policy for rare disease and initiated the European Project for Rare Diseases National Plans Development (EUROPLAN) in 2008. In order to provide recommendations on strategies and plans for rare diseases on the national and regional levels for each Member State, EUROPLAN developed a set of indicators that were divided into six categories according to their intended measure. These include a) health outcomes, b) risk factors, c) intervention coverage, d) structure, e) process, and f) non-health related results (13).

In 2009, the EU Council issued a call for Member States to develop and adopt national plans or strategies to guide the implementation of actions specifically in the field of rare diseases (14). From 2009 to 2013, the EU Committee of Experts on Rare Diseases (EUCERD) was commissioned under the Seventh Framework Programme for Research and Technological Development (FP7) to monitor, recommend actions and improvements, and strengthen European and international ties in the field of rare diseases.

The importance of establishing an international strategy for rare diseases should not be underestimated. Better coordination among funding agencies and between research groups, clinicians and patients is paramount to address gaps in scientific knowledge and advance research in rare diseases. Research networks bring together experts from different fields of work (e.g. scientists and specialists with extensive clinical experience) and from different sector (e.g. the public and private sectors).

A wide variety of European projects have focused on delivery of care for rare diseases, funded by EU FP7 grants. In PCD, the Better Experimental Screening and Treatment for Primary Ciliary Dyskinesia (BESTCilia) programme characterised the natural history of the disease and improved

diagnostics and treatment for patients across Europe through vast collaboration between clinicians and researchers from 12 European and North-American countries (4, 15-19). Better Experimental Approaches to Treat Primary Ciliary Dyskinesia (BEAT-PCD) (http://www.beatpcd.org/) was an international network of experts from diverse clinical (e.g. paediatric and adult pulmonology, ear, nose and throat specialist (ENT), physiotherapy) and scientific specialties (e.g. genetics, imaging, cell biology, microbiology, bioinformatics), working with industrial partners. BEAT-PCD was supported by EU- Framework Horizon 2020 (COST Action BM1407) and aimed to identify gaps in knowledge and facilitate PCD-related collaborative research to identify mechanisms, study disease patterns and progression, define outcome measures, improve clinical management, and identify high priority therapies (71,131) .

2.1.3 Specialised Centres of Reference

Improvement of patient outcomes depends on clearly defined care pathways that enable efficient evidence-based diagnosis and patient care. Whilst true for all diseases, it is more evident in rare diseases as local and national resources are often scarce and therefore coordination and resource allocation are crucial. In 2017, the European Reference Network Board of Member States approved a new initiative in the field of rare diseases through the formation of 23 European Reference Networks (ERNs). These networks were formed through collaboration between centres of expertise (i.e. specialised centres) and healthcare providers across Member States. The aim was to facilitate patient access to highly specialised clinical care for rare diseases, where a concentration of resources and expertise is available. Prior to this initiative, only border regions were covered by bilateral healthcare cooperation agreements between involved European countries; establishment of ERNs will guarantee patients' rights for cross-border healthcare from within the entire territory of all Member States (20).

PCD is one of the nine core networks of the ERN on Rare Respiratory Diseases (ERN-LUNG). The aims of ERN-LUNG include addressing many of the challenges discussed in 1.1.1, such as increasing referrals of patients by enabling cross-border access to specialist centres; developing standardised management guidelines and equal access to treatments; supporting the development and maintenance of disease-specific registries; promoting a multidisciplinary approach to patients care; and providing a platform for exchange of knowledge between experts and for communication with patients (14).

2.2 Primary ciliary dyskinesia

PCD is a rare heterogeneous genetic disease characterised by abnormal ciliary function. Motile cilia lining the respiratory epithelia from the nasal cavity to the terminal bronchioles, the Eustachian tubes, and sinuses beat in a synchronised way to clear mucus, bacteria and debris. Mucociliary clearance is an effective defence mechanism in the airways against bacteria and other toxic substances.

Ciliated epithelium can also be found in the Fallopian tubes and therefore cilia movement assists in transporting the egg cells from the ovaries to the uterus. Motile cilia share a common axonemal structure with the spermatozoa flagella. Cilia are also present in the brain ependymal, driving cerebrospinal fluid to the cerebral ventricles.

2.2.1 Cilia ultrastructure and function

The normal ultrastructure arrangement of motile cilia consists of nine microtubule doublets and a single central pair, the “9+2” formation (Figure 1). Dynein arms are attached to the doublets and are responsible for the mechanical action of the cilia. The outer dynein arm (ODA) is the main contributor to cilia movement and is responsible for initialising and generating the force of the beating stroke. The inner dynein arm (IDA) is linked to the control and amplitude of axonemal bending (21). Nexin proteins connect the doublets to each other and radial spoke projections from the dynein regulatory complex (DRC) toward the central pair maintain the orientation of the microtubules. The central pair controls the forward and backward planar motion of the ciliary beating (21).

The embryonic node exhibits motile cilia with a “9+0” formation, lacking the central pair. As a result, cilia beat in a circular manner and the flow generated by this determines the left-right pattern of major organs during embryogenesis. Non-motile cilia (i.e. sensory cilia), found in the inner ear, eye, renal tubules, bile and pancreatic duct, bone, cartilage and fibroblasts, also have a similar “9+0” assembly. This explains the occasional overlap between PCD and other rare ciliopathies such as Bardet-Biedl, Joubert, Meckel-Gruber and Alstrom syndromes, orofacialdigital syndrome, retinitis pigmentosa and autosomal recessive polycystic kidney disease (22).

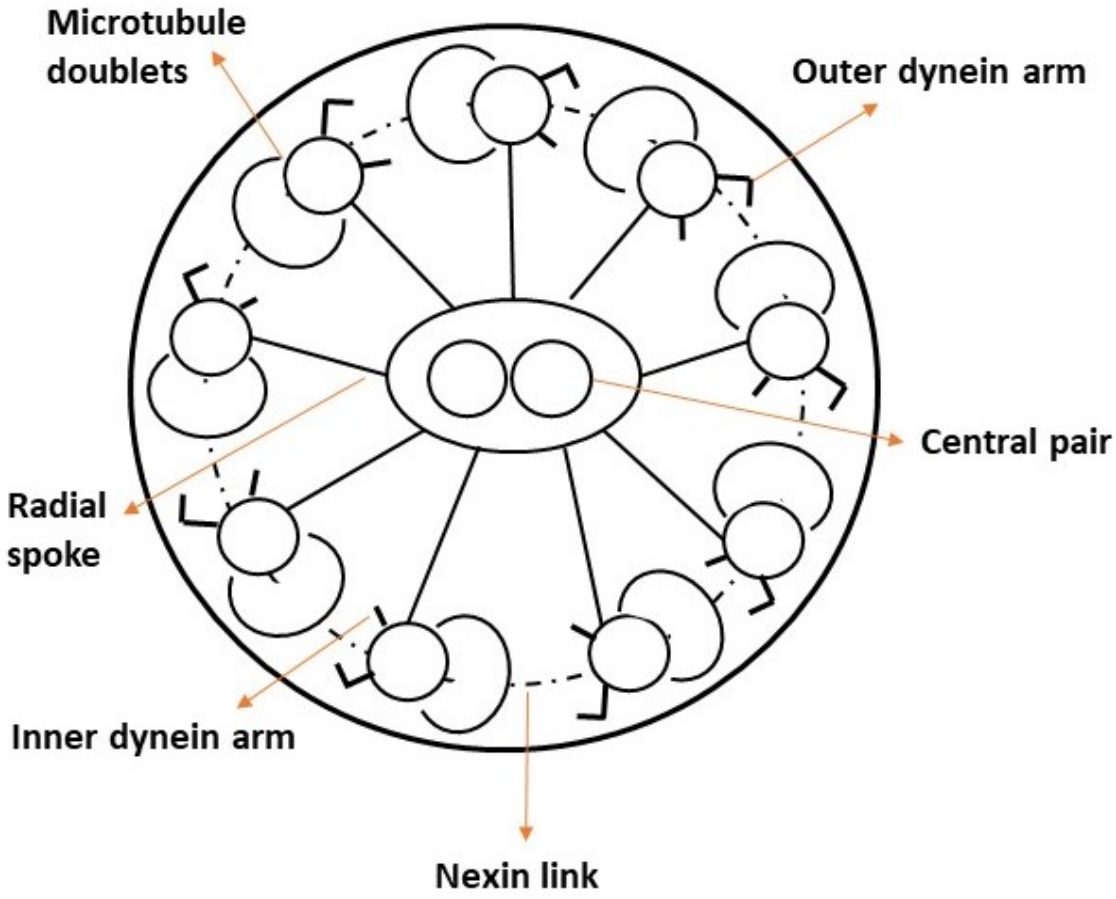


Figure 1. Schematic depicting normal ultrastructure of a cross-section of motile cilium.

An effective ciliary beating in motile "9+2" cilia is composed of a unidirectional, coordinated stroke and a recover phase. The cilium moves in the direction of the flow generated during the stroke phase, while in recovery the tip of the cilium slowly moves backwards. Ciliary length and beat frequency determine the speed of the flow produced, which can go up to about one millimetre per second (23).

2.2.2 PCD genetics

PCD is caused by mutations in genes that encode ciliary proteins, and proteins involved in assembling, transporting and docking ciliary components. It is usually inherited as an autosomal recessive condition, requiring mutations to be identified in both alleles. However, X-linked mutations in *RPGR*, *OFD1* and, more recently, *PIH1D3* have been associated with PCD (24, 25). Mutations have been identified over 45 genes to date, but more remain to be discovered as pathogenic variants cannot be identified in up to 30% of patients with observed ultrastructural defects by transmission electron microscopy (TEM) (26-29). Cilia have a complex structure and mutations in any of the genes encoding one of the numerous proteins involved in ciliogenesis, cilia regulation and cilia function could potentially lead to cilia impairment or immotility. Unlike

most genetic diseases, the majority of mutations that cause PCD are private or unique to families, with only a small number of founder (or recurring) mutations, isolated from specific populations such as patients with Hispanic descent, the Amish, the Bedoin, and the South Asian communities in the UK (30-35). The large number of genes and the even greater number of mutations are expressed as a clinically and genetically heterogeneous disease with a range of clinical phenotypes.

Approximately 85% of mutations in PCD are nonsense and deletion mutations leading to truncated protein or loss of function, independent of the gene mutated (36, 37). A nonsense mutation is a point mutation resulting in a stop codon, with the resulting protein being incomplete, truncated or non-functional. Deletion mutations consist of loss of a sequence of DNA, which can vary from small sections to entire pieces of chromosome. When deletions occur in sequences of DNA that contain a number of bases not divisible by three, the translation process is forced to shift its frame, consisting of what is called 'frameshift' mutation. Deletions can also occur in codons (i.e. groups of three DNA bases), leading to 'in-frame' deletions. Missense mutations are also found in PCD but cannot be as easily attributed to causing the disease, as they translate to the loss of a single amino acid in the protein chain. These mutations must be seen in both parents and have a compatible ultrastructural defect seen in TEM to be considered pathogenic.

The genes that encode the ODA components include those involved in the heavy chains (i.e. *DNAH5*, *DNAH11*), intermediate chains (i.e. *DNAI1*, *DNAI2*), and light chains (i.e. *DNAL1*, *NME8*). The attachment of the ODAs to the outer doublets is mediated by genes that encode the ODA docking complexes (i.e. *CCDC114*, *ARMC4*, *CCDC151* and *TTC25*). Genes involved in the cytoplasmatic pre-assemble of dynein arm components result in ODA and IDA defects. These include *DNAAF1*, *DNAAF2*, *DNAAF3*, *DNAAF4*, *DNAAF5*, *DNAAF6*, *LRCC6*, *ZMYND10*, *SPAG1*, *C21ORF59*, *CFAP300* (26).

2.2.3 Clinical manifestations

Abnormal function of epithelial cilia in lungs, nasal and sinus cavities, and ears leads to recurrent and chronic infections of upper and lower airways (5). Over 60% of full-term neonates suffer from unexplained neonatal respiratory distress syndrome (NRDS) (17, 38, 39), which typically occurs 12 to 24 hours after birth. Neonates are often admitted to intensive care units with NRDS or neonatal pneumonia and require oxygen therapy due to hypoxaemia, with atelectasis or lobar collapse often seen on chest radiography. Persistent daily wet cough is one of the most frequently reported symptoms throughout life (17, 40). Chronic suppurative pulmonary infections are

frequent and progressive, leading to irreversible lung damage and bronchiectasis, which can be found in up to 22% of children at diagnosis and almost universally in adults (17, 38, 41, 42).

Rhinitis can be present in approximately 25% of neonates with PCD and can persist throughout life, with rhinorrhoea reported in more than 90% of adults with PCD (17, 42). Older children and adults often suffer from chronic sinus disease (17, 40, 43). Chronic rhinosinusitis, defined as an inflammatory disease of the paranasal sinuses with symptoms lasting over 12 weeks, is reported in over 70% of children with PCD (10, 44, 45). Symptoms include facial pain, persistent nasal secretions or blockage, and decreased sense of smell. Opacification of the paranasal sinuses can be observed in CT scans of almost 90% of adults with PCD and hypoplasia or aplasia of the frontal or sphenoid sinuses in 73% of patients with PCD and 36% of adults in particular (42, 46). Nasal polyps can be observed in approximately 30% of patients (42).

Recurrent acute otitis media (OM) is present in up to 80% of patients at an early age (10). Symptoms seem to decrease significantly with age, with 68% of children under the age of six reporting three or more episodes in a period of six months or at least four in the last 12 months, but seen in only 17.6% of those aged between 12 and 17 years (47, 48). Chronic serous OM with effusion (glue ear) is present in over 90% of children up to the age of 17 and is often persistent, lasting at least three months (48). However, less than a quarter of adults have chronic OM. Conductive hearing loss (≥25 decibels) has been reported in over 60% of children under the age of five, which is particularly concerning as it can lead to speech impairment (23). Hearing loss decreases with age but is still high in adults (42, 48).

S*itus inversus totalis*, defined as mirror image arrangement of all organs, can be found in approximately 50% of cases due to disruption of embryonic nodal flow (24–26). Therefore, mutations on genes that result in lack of central pair are not associated with organ laterality defects (49). 12.1% of patients with PCD present some degree of heterotaxy (or situs ambiguous), which is strongly correlated with the presence of congenital heart defect (CHD). CHD can be present in up to 17.1% of children (27) depending on the classification system adopted (49).

The sperm flagella and the fimbriae of fallopian tubes exhibit a similar structure to respiratory cilia, with studies reporting fertility problems in up to 50% of men and in varying degree in women (40, 50-53). However, data are scarce and there are very few recent studies investigating fertility problems in patients with PCD.

A meta-analysis found variations on the reporting of clinical manifestations in PCD (40). Only a third of included studies reported on NRDS, and only a quarter on CHD. It is unclear if these findings reflect limitations in data collection by the authors, issues in recording and reporting

these symptoms by clinicians, or a combination of both. Symptoms described in most studies lacked standardisation, limiting comparisons between them. For example, 14% of the studies that reported situs abnormalities only provided information on cardiac situs or used the terminology '*situs ambiguous*' without defining the term. The need for standardised data led to development of FOLLOW-PCD, a disease-specific follow-up data collection form for prospective use in clinical practice and research (54).

2.3 Referral for diagnostic testing

Clinical presentation of PCD is heterogeneous and patients exhibit different combinations of symptoms, some of which can vary over time (16,17). Moreover, the symptoms are not disease-specific; therefore, patients need to be correctly identified by general clinicians, paediatricians, obstetricians, ENT specialists, pulmonologists, fertility specialists, neonatologists and cardiologists. Although not specific in isolation, the pattern, combination and chronology of symptoms can be strongly suggestive (18,19).

Lack of awareness of PCD by general practitioners and paediatricians was reflected in an international survey of 271 PCD patients. They reported that 37% of PCD patients had over 40 visits to medical professionals due to PCD-related symptoms before being referred for testing, a considerably high diagnostic delay interval (45). The reported prevalence of PCD in Switzerland was 1:63 000, which represents 0.16:10 000, indicating that many patients remain undiagnosed (55).

In older patients, a radiological finding of bronchiectasis due to unknown cause is a common reason for referral of patients to diagnostic services, with reports of up to 13% of patients with diffuse bronchiectasis due to PCD (56). Studies have shown that bronchiectasis can be detected in 50% of children over 8 years-old (48) and up to 98% of adult PCD patients (38). It is probable that most undiagnosed cases are currently being managed as idiopathic bronchiectasis, which is problematic because PCD is a multisystem disease and therefore requires specialist management by a multidisciplinary team.

The development of screening tests and predictive tools is paramount to improve accurate identification of patients for diagnostic testing. Simple and easy-to-use predictive tools that provide disease probability scores, such as the PrImary CiliARy DyskinesiA Rule (PICADAR) and the expert-defined clinical symptoms scores, will likely improve referrals from secondary care, but require validation in different settings (17, 57, 58). PICADAR consists of a predictor model that includes seven variables: full-term gestation, neonatal chest symptoms, admission to neonatal intensive care, chronic rhinitis, ear symptoms, *situs inversus* and CHD. It was shown to be accurate

for patients with persistent wet cough using a cut-off score of 5 points, with a sensitivity and specificity of 90% and 75% respectively in a retrospective validation study of 187 patients consecutively referred for diagnostic testing (17). The expert-led clinical symptom score includes four variables: unexplained neonatal respiratory distress, early-onset year-round wet cough, early-onset year-round nasal congestion, and laterality defect (57). The authors report sensitivity of 80% and specificity of 72% when two variables are present.

Despite the majority of patients presenting with neonatal respiratory symptoms (39) and up to 84% of patients having a history of unexplained oxygen requirement after birth, studies have shown that neonates are rarely diagnosed with PCD nor referred for testing (4, 11, 59). Mullowney *et al* (39) demonstrated that the combination of lobar collapse, *situs inversus totalis*, and/or a persistent oxygen therapy for over 2 days in term newborns with respiratory distress can accurately predict PCD. The diagnosis should also be considered in term neonates with normal situs. Equally, having situs abnormalities does not equate with PCD. Only a quarter of individuals with *situs inversus* referred for diagnostic testing were diagnosed with PCD in a recently published UK study (60).

The European Respiratory Society (ERS) guidelines for the diagnosis of PCD provides recommendations on patients that should be referred to centres that can provide diagnostic testing (25). The following combination of symptoms should trigger the referral process: a) persistent wet cough, b) situs abnormalities, c) CHD, d) persistent rhinitis, e) chronic OM with or without hearing loss, f) history of upper or lower respiratory symptoms in term neonates, or g) admittance to neonatal intensive care unit. The guidelines highlight that patients without situs abnormalities should be referred for testing if they have symptoms suggestive of PCD. Additionally, patients with a positive family history, particularly if coupled with clinical symptoms, should be tested for PCD. The guidelines also recommend the use of predictive tools such as PICADAR when deciding who to refer for diagnostic testing (17).

A Europe-wide survey found that differences in referral patterns influence the disease prevalence reported (4). The active search for potential PCD cases in the only major hospital in Cyprus, where the paediatric pulmonologist has an interest in PCD, led to a higher than average disease prevalence. A retrospective evaluation of data from ten years of experience from the Southampton PCD Diagnostic Service found that 8% of patients referred to the centre were diagnosed with PCD (61). However, there have been no studies comparing the effectiveness of different referral pathways and their impact on access to services.

Most patients were willing to travel longer distances to have access to specialised tests and care (11). Centres managing over twenty PCD patients were associated with diagnosis at an earlier age,

highlighting the benefit of high throughput of cases for training of clinicians, healthcare professionals and diagnostic scientists (4). Presence of situs abnormalities, a rare condition in the general population, has an impact on age at diagnosis, with children presenting with *situs inversus totalis* diagnosed at an approximate median of 4 years of age whereas those with *situs solitus* (or normal organ arrangement) diagnosed significantly later (approximately 6 years of age) (4, 41, 62). Recent data from England revealed a decrease in the median age at diagnosis to 2.6 years in 2015, significantly lower for those with situs abnormalities (1 vs 6 years for *situs solitus*) (63).

However, deciding when and where to refer patients is not always straightforward. A variety of factors can influence referrals, such as individual initiative by medical professionals, awareness of the existence of diagnostic centres, and national and local funding structures. Centralisation of care will likely provide access for patients to expert centres using standardised protocols with quality assurance *e.g.* ERN-LUNG (64) and the North American Genetic Disorders of Mucociliary Clearance Consortium (GDMCC).

2.4 PCD diagnostic testing

There is no single 'gold standard' test to diagnose PCD; instead, diagnosis is reached through a combination of diagnostic tests coupled with a suggestive clinical history, advocated by both the ERS and the American Thoracic Society (ATS) evidence-based diagnostic guidelines (25, 65). Patients with clinical symptoms suggestive of PCD should undergo multiple testing, which can include nasal nitric oxide (nNO) assessment, HSVA and epithelial cell culture, TEM, genetic testing, immunofluorescence and electron microscopy tomography.

2.4.1 Nasal nitric oxide

Levels of nNO are significantly lower in patients with PCD compared to healthy individuals and those with other respiratory diseases such as cystic fibrosis (CF), with a sensitivity ranging from 90 to 100% and specificity from 75 to 97% (66) depending on the adopted cut off and reference standard. Healthy individuals usually have nNO levels above 300 $nL \cdot min^{-1}$, while most PCD patients have levels below 77 $nL \cdot min^{-1}$ (67). Equally, nNO levels in PCD patients are discriminatorily lower than in those with other respiratory illnesses (e.g. asthma, allergic rhinitis, chronic obstructive pulmonary disease, non-CF bronchiectasis), chronic sinusitis without nasal polyposis and immunodeficiency (67-71). However, the precise mechanism through which nNO is lower in PCD remains unknown (72, 73).

Ideally, a technically acceptable plateau should be obtained using a stationary chemiluminescense analyser, preferably during a velum closure manoeuvre to reduce contamination from lower

airway by isolating the nasal cavity (*e.g.* exhalation against resistance or breath hold), as levels of NO in the lungs are ten to 100 times lower than in the nasal cavities (72).

Velum closure manoeuvres are recommended for adults and children over five years of age. They are less reliable in younger children as they might not be able to perform the test since the manoeuvres requires significant cooperation from the patient. Additionally, levels of nNO were shown to be lower in healthy young children, likely due to under-developed sinus cavities.

Recent studies have investigated nNO levels during tidal breathing for infants (74). Marthin *et al* (75) showed that children with PCD up to the age of two years had low levels of nNO measured during tidal breathing, similar to levels seen in healthy infants during respiratory infections. However, low levels in healthy infants increased on follow-up measurements after resolution of infection and with age whereas they remained low in patients with PCD. Since nNO is measured in patients suspected of having PCD when they are infection-free, low levels can be discriminatory. Importantly, the study showed a nearly 100% success of tidal breathing in this age-range, demonstrating that young children can reliably perform the technique. However, there is still a lack of standardised references for tidal breathing and different age-ranges, and study data were based on historical data from only seven patients, limiting conclusions (76).

The North American technical paper on standardising nNO measurements recommends the use of nNO in older children (≥ five years old) and adults with a high pre-test probability of PCD, as diagnostic accuracy of the test is considerably lower when key clinical features are not present. Children between the ages of two and five years can perform tidal breathing manoeuvre, but results should not be taken in isolation and nNO measurements need to be repeated after the age of five. The technical paper does not recommend testing nNO levels in those younger than two years of age (72).

Collins *et al* (77) reported a positive predictive value (PPV) of 42.6% from 282 consecutive referrals, of which 11% were PCD positive patients. The North American guidelines strongly recommend repeat testing on separate visits at least two weeks apart using velum closure techniques if levels are lower than reference to ensure they are not a result of secondary factors such as acute viral infection (65). After exclusion of CF, false positives can occur due to poor technique, severe nasal obstruction, blood in the nasal cavity, recent sinonasal surgery, naso-oro-facial malformations, and obstruction of the sampling line. False negatives can occur when ambient NO is high and in PCD patients with specific gene mutations.

The ERS guidelines argued that nNO should not be used in isolation or as a general screening test, as it would perform poorly in populations with low prevalence of PCD (*e.g.* primary and secondary

care settings). Instead, it should contribute to the final diagnostic decision (25, 77). In contrast, the ATS guidelines recommended the use of nNO as a confirmatory test when performed according to standardised protocols in a population with high probability of having PCD and after excluding CF (65, 67, 72). However, nNO alone should also not exclude a diagnosis of PCD as normal levels have been described in patients with mutations in *CCDC103* and *RSPH1* (35, 78).

Reporting of nNO measurements should be converted from parts per billion (ppb) to nanolitres per minute ($nL \cdot min^{-1}$) by multiplying ppb levels by the sample gas flow rate of the analyser. This will allow for comparisons between measurements performed using devices with different flow rates.

2.4.2 High-speed video analysis

HSVA is the only diagnostic test routinely used in clinical practice that directly assesses ciliary function. Ciliated respiratory epithelia are obtained by brushing or biopsy of the nasal mucosa using a cytology brush. Strips of ciliated cells are suspended in buffered culture medium and recorded using a light high magnification microscope fitted with a high-speed video camera. The images are then played back at slow speed to allow visualisation of cilia beating. As presence of ciliated cells and coordination of cilia movement can be affected by recent infection (i.e. secondary ciliary dyskinesia), patients should be free of infection for 4 to 6 weeks prior to sampling (79). These ciliated samples are used for HSVA, TEM and any additional investigations (e.g. immunofluorescence and cell culture).

Ciliary beat frequency is calculated based on the time required for a group of cilia to complete ten ciliary beat cycles, usually calculated manually. Respiratory cilia typically beat at a frequency of 8 to 20 Hz; however, ciliary beat frequency (CBF) is dependent on temperature, and mechanical and chemical stimulation, such as osmosity, viscosity, oxygen concentration, humidity, and extracellular and intracellular pH. (60, 80) When used in isolation, CBF has poor predictive value to diagnose PCD as patients with central complex defects have a beat frequency within the normal range (21). Therefore, recent guidelines advocate evaluation of ciliary beat pattern (CBP) as a more reliable parameter to guide diagnosis. This consists of detailed observation of the movement of cilia, in order to determine if the pattern seen consistently across a set number of cell clusters corresponds to what would be expected as effective stroke and recovery phases with mucociliary clearance, as described in 2.2.1. Typical findings in PCD patients include immotile cilia, stiff cilia with incomplete forward and recover stroke, or rotating cilia observed from a top view.

Recent studies have shown associations between specific changes in CBP, hallmark ultrastructural defects, and genotype. For example, static cilia are seen in patients with combined absence of

outer dynein arms (ODA) and inner dynein arms (IDA) (*e.g. ZMYND10*), while ODA defects alone (*e.g. DNAH5*) are associated with residual movement and vast areas of static cilia. Circular beating from top view is observed in those with central complex defects (21, 81).

Typical CBP cases are relatively easy to detect by HSVA; however, subtle changes are difficult to differentiate from secondary defects and are only likely to be picked up by specialised scientists in high throughput clinics, and, even then, risk being missed. Analysis by scientists with substantial experience in HSVA is therefore essential (82).

2.4.3 Transmission electron microscopy

TEM is used to visualise respiratory cilia ultrastructure in transverse sections of chemically fixed epithelium. TEM was once the 'gold standard' for PCD diagnosis (83); however, it is now known that 15 to 20% of PCD cases have disease-causative mutations but normal ultrastructure as seen by TEM (e.g. *DNAH11, HYDIN*) (36, 84). Additionally, the ultrastructure of cilia can be affected by secondary defects due to infection, inflammation or sample processing (85). Appropriate collection, processing and interpretation of samples by experienced scientists is therefore essential.

The recently published TEM consensus statement, developed in collaboration with experts in electron microscopy from 18 PCD diagnostic centres in 14 countries, provided detailed information on how to classify each ciliary defect (85). Importantly, the tested and externally validated consensus emphasised the need for clear and standardised diagnostic reports containing information on the source and adequacy of the sample, number of cross sections assessed, percentage of abnormal cilia, consistency of defect across several cells, and succinct summary of key findings, including class defects.

Absence of ODA, combined absence of ODA and IDA, or absence of IDA with microtubular disarrangement (MTD) in the majority of cilia cross sections are considered class 1 defects (*i.e.* hallmark) and when observed confirm a diagnosis of PCD (25, 85). Central complex defect, mislocalisation of basal bodies with few or no cilia, isolated MTD, ODA defects in 25% to 50% cross sections, or combined IDA and ODA defects in 25% to 50% cross sections, are classified as class 2 defects (*i.e.* non-hallmark defects). They indicate a diagnosis of PCD when consistently seen across more than one sample or after cell culture, in patients with suggestive clinical symptoms, and combined with consistent results from other diagnostic tests.

Sample adequacy needed to classify a class 1 defect was defined as assessment of at least 50 axonemes in transverse sections with an intact ciliary membrane from several cells. Since the

ultrastructure of cilium at the tip is different from the base, cross sections should be taken from both the proximal and distal regions of the axoneme.

2.4.4 Genetic testing

Pathogenic bi-allelic mutations in over 45 genes have been published to date (26). These are estimated to be responsible for approximately 70% of PCD cases, and as new genes are discovered this is likely to improve (28, 29, 85). Candidate gene analysis enabled the discovery of the first gene attributed to PCD (i.e. *DNAI1*) and is still the main tool used for gene identification, but next-generation sequencing technologies might offer a greater insight once technical challenges are overcome (32, 86).

While genetic testing continues to evolve, its role in the diagnostic pathway is still unclear. Most studies to date focused on the discovery of novel genes and were conducted in small cohorts with PCD already confirmed by TEM; therefore, the accuracy of genetic testing remains unknown. However, genetic testing can provide a diagnosis in cases where HSVA and TEM are normal or subtly abnormal (*e.g. DNAH11* and *CCDC103*) (35, 36).

Genetic testing has significant limitations. Test accuracy depends on gene panel used, with improved sensitivity as the number of tested PCD genes increases. For example, sensitivity increased from 71.9% when testing 12 genes to 93.9% when testing 32 genes with deletion or duplication analysis in patients with a clinical diagnosis of PCD (29, 65).

2.4.5 Other diagnostic tests

HVSA can be performed after culturing of epithelial cells, using either the air-liquid interface (ALI) or the spheroid culture methods, to identify alterations in cilia movement. Culturing epithelial cells in a controlled environment, with the use of antibiotics, can help differentiate primary from secondary defects as infection-derived issues will likely be prevented. Abnormalities can be attributed to primary defects of cilia function if the abnormal movement, seen by HSVA of the primary sample, persists in the cultured sample (87).

Immunofluorescence is based on the visualisation of fluorescence-labelled antibodies that bind to specific cilia proteins in epithelial cells visualised by a fluorescent or confocal microscope. Antibodies targeting ODA, IDA, radial spoke head and DRC proteins are available but do not cover all known ultrastructural defects. The majority of published studies assessed the implications of different genetic mutations on cilia protein production and were therefore conducted on patients with an established PCD diagnosis (25). Shoemark *et al* (88) recently published the first diagnostic

study assessing test accuracy using a panel of six antibodies. They showed that immunofluorescence correctly identified 22 out of 25 PCD patients and 100% of 252 consecutive patients in whom a PCD diagnosis was considered highly unlikely. These findings strongly suggest that immunofluorescence can play an important role in the diagnostic pathway but there is need for a wider range of antibodies, with improved quality, and for standardisation of administration and reporting of tests.

Electron microscopy tomography is a computer-assisted TEM technique that allows for generation of three-dimension images, providing high-resolution projections of cilia ultrastructure. There have been no diagnostic studies to date but Shoemark *et al* (89) have shown that patients with mutations in *HYDIN* or *DNAH11*, which appear to have normal ultrastructure using standard TEM, in fact have abnormal ultrastructure when seen using electron microscopy tomography (36, 90).

2.4.6 Diagnostic guidelines

The ERS guidelines provides evidence-based recommendations on PCD diagnosis and offers an algorithm for diagnostic testing, which should be based on clinical symptoms, nNO assessment, HSVA and epithelial cell culture, TEM, and genetic testing (25). Other tests, such as immunofluorescence of ciliary proteins and electron microscopy tomography are currently available in only a few PCD centres but might offer valuable information, particularly in difficult diagnostic cases.

Additionally, the guidelines provide a classification system to determine diagnostic certainty based on results available from the tests performed (Figure 2) (25). According to the ERS diagnostic guidelines, patients with a hallmark ciliary ultrastructural defect by TEM (*i.e.* combined ODA and IDA defects, isolated ODA defects, or IDA defect with MTD) and/or bi-allelic or X-linked mutations in known PCD genes are classified as 'PCD positive' (85). 'Highly likely PCD' is assigned to patients with compatible history, very low nNO and either a) typical and consistent findings on HSVA (e.g. immotile or circular beating cilia) on three separate occasions or b) suggestive findings on HSVA following cell culture. No single test can exclude the diagnosis but in patients presenting with modest or low clinical suspicion and normal or high nNO levels with normal HSVA, the diagnosis is deemed 'extremely unlikely PCD' and no further test is usually warranted. Patients with either 'definitive positive' or 'highly likely PCD' diagnosis should receive treatment for PCD. Asserting the diagnostic certainty is an important step to ensure correct enrolment of confirmed cases in prospective cohorts and RCTs.

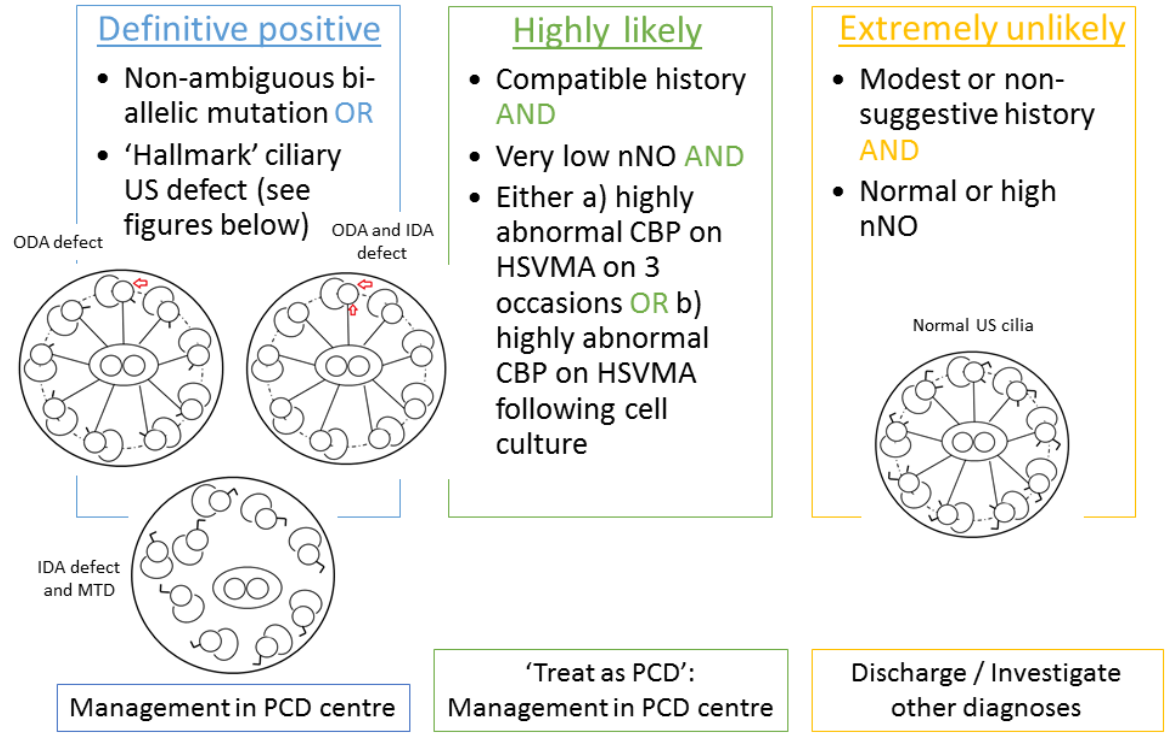


Figure 2. PCD classification according to the ERS guidelines. Figure created by me, previously published in the European Respiratory Review (91).

The ATS clinical practice guidelines for the diagnosis of PCD in children and adults also provided evidence-based recommendations (65). A panel of PCD specialists summarised and reviewed the evidence to recommend for or against the substitution of the reference standard, which consisted of hallmark TEM or genetic testing for ≤12 genes, by four potential diagnostic tests: extended genetic panel with over 12 genes, nNO measurements, HSVA, and CBF. The panel concluded that a diagnosis can be confirmed by the presence of biallelic mutations in PCD genes, TEM hallmark defect or low nNO levels, defined as values below 77 $nL \cdot min^{-1}$ on two separate occasions measured by a chemiluminescence device using standardised protocol at a PCD specialist centre (72). They recommended against using CBP by HSVA as a substitute to the reference standard. The proposed diagnostic algorithm includes nNO testing for all patients older than five years of age with at least two of the four suggestive clinical features according to Leigh *et al* (57). If nNO is not available or not possible (*e.g.* younger children), extended genetic testing panel should be conducted, followed by TEM where a biallelic pathogenic variant in a PCD-causative gene cannot be identified. If TEM is normal, a PCD diagnosis should still be considered possible.

Differences between the recommendations made by the ERS and the ATS guidelines might reflect variations on definition of reference standard for comparison to index test (*i.e.* as a substitute in case of the ATS, or as a confirmatory test or to be used in conjunction with other tests in case of the ERS) (92). There was an additional year between the two publications, and therefore the ATS

included studies that were not available when the ERS guidelines were being developed. Additionally, the grading classification of evidence by the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) criteria also changed between 2017 (ERS guidelines) and 2018 (ATS guidelines). Despite these differences, both concluded that hallmark TEM and biallelic mutations in know PCD genes confirm a diagnosis and that there is no stand-alone test or combinations of tests that can exclude PCD cases with 100% certainty.

The availability of diagnostic tests varies between countries and it might not be feasible to reach a definite diagnosis in centres with few resources due to financial limitations and lack of technical expertise (4, 93). Over 20% of patients included in the European PCD registry had a diagnosis based only on symptoms, and a systematic review of clinical symptoms found that 10% of studies relied on clinical diagnosis alone (40, 94).

2.4.7 Standardised reporting for diagnostic testing

While advances have been made to standardise the diagnostic pathway, improvements are now needed on the processing, analysing and reporting of results from diagnostic tests. Cut-off values often vary considerably between centres, as they depend on local expertise, equipment, techniques, and laboratory and sample conditions (95).

Levels of nNO were found to be significantly lower in PCD patients compared to healthy and disease controls in a recent meta-analysis (66). However, emphasis was given to the high levels of heterogeneity seen between studies (I^2 = 93.9% when comparing PCD to healthy individuals and 40.8% when comparing to CF patients), reflecting differences in study design, age, control group and breathing techniques. Only four studies reported on nNO levels for consecutive referrals, of which half lacked basic information on sampling and patient diagnosis. Equally important is the adoption of appropriate cut-off values that can be applied to different settings, analysers and populations. Leigh *et al* (67) have shown that 77 $nL \cdot min^{-1}$ is a disease-specific discriminatory cut-off value reproducible in six other North-American sites when adopting a standardised protocol. However, there is still no consensus on appropriate thresholds for each technique, particularly for children less than 5 years of age in whom tidal breathing is performed (74-76).

Researchers have questioned the validity of CBP in qualitatively describing cilia impairment through HSVA due to the observer-dependent nature of the test (96). Beat pattern abnormalities can also occur secondary to infection, damage during sampling or inflammation of the epithelium (97). There is lack of standardisation in definitions of CBP, with scientists applying different categories for similar abnormalities. This lack of consistency has hindered broader comparisons and pooling of data from different countries. Quantitative analysis of pattern parameters, such as

weighted distance travelled per second and sheer stress, might assist detection of abnormalities, but have not yet been evaluated in complex cases (98, 99).

A quantitative approach to cilia evaluation is highly recommended when reporting TEM findings, as intermittent ultrastructural defects can occur (100). The recently published international consensus guidelines provides clear guidance on standardised analysis and reporting of TEM results for the diagnosis of PCD (see 2.4.3) (85). However, guidelines on standardised methodology for sample processing and visualisation are still needed, as differences in equipment and local procedures remain.

Reporting of genetic testing results must follow international guidelines. Segregation of variants from parents should be performed, as benign variants and variants of unknown significance are regularly identified because the genes involved in PCD are large (101). Therefore, the specific mutations need be observed in both patient and parents in order to establish a genetic diagnosis. Additionally, the pathogenicity score must be taken into consideration as variants of unknown significance might not be pathogenic. Consistency between genotype and other diagnostic test results such as TEM, HSVA and immunofluorescence is also important and should be checked before reaching a diagnosis.

2.5 PCD management

Currently there is no cure for PCD and therefore management of patients aims to improve quality of life and psycho-social wellbeing, and to prevent progression of irreversible lung disease. Like with other rare diseases, the evidence-base for treating patients with PCD is poor. The choice of management strategy is based on expert consensus derived from other diseases such as CF and non-CF bronchiectasis, and the clinician's personal experience and preference.

There are two expert consensuses on PCD management: the ERS task force consensus in children published in 2009 (102), and the North American PCD Foundation consensus published in 2016 (103), neither of which are evidence-based. Both recommend regular follow-up visits at least every six months, conducted in specialist centres with MDT team experienced in managing chronic suppurative lung disease. In the case of management of children with PCD, specialist teams should include pulmonary paediatricians, ENTs and respiratory physiotherapists.

Baseline assessments should be conducted soon after diagnosis. These include spirometry measurements, audiometry, sputum cultures and chest imaging. Additional investigations such as cardiac and abdominal screening should be considered.

2.5.1 Lower airway management

Respiratory management is largely extrapolated from CF and non-CF-bronchiectasis despite considerable differences between PCD and CF pathophysiology, prognosis and response to treatment (104, 105). The cornerstones of lower airway management are regular monitoring of lung function and respiratory symptoms, respiratory physiotherapy coupled with exercise, and active treatment of infections with antibiotics. As in all chronic respiratory conditions, annual influenza vaccine and routine immunisations are recommended (102, 103).

2.5.1.1 Treatment

Airway clearance through physiotherapy and exercise is strongly recommended and should be initiated as soon as the diagnosis of PCD is considered likely. There is no evidence or clear consensus on the most effective physiotherapy technique, with practices varying between centres. Disease severity, and patient age, preference and ability in performing airway clearance techniques are factors to be considered when selecting the most appropriate physiotherapy method. In England, data from annual reviews conducted in 2015 showed that most children were using either oscillatory positive expiratory pressure or positive expiratory pressure masks (6). Physical exercise should not be a substitute to respiratory physiotherapy but rather an adjunct method to assist in airway clearance in patients with PCD.

Osmotic agents that can promote mucociliary clearance are important to facilitate movement of secretions, reduce infections, and ultimately prevent progressive lung damage. Nebulised hypertonic saline is used as an adjunct to airway clearance in PCD and is a relatively cheap therapy. A recently published double-blind crossover RCT did not show any clinically significant difference in health-related quality of life (HRQoL), measured by the St George's Respiratory Questionnaire, between patients on 7% hypertonic saline and isotonic saline (106). However, it was underpowered, the primary outcome was not disease-specific, and the control group received nebulised isotonic saline, which might have some therapeutic benefit in PCD (107). It is routinely used in England, with half of the children under management at PCD specialist centres prescribed hypertonic saline to aid airway clearance (6).

Use of inhaled β-agonists should be restricted to cases with proven airway reversibility as there is no evidence to suggest that regular use of bronchodilators is beneficial in PCD (108). Recombinant human deoxyribonuclease I (rhDNase) promotes clearance of secretions by reducing mucus viscosity in the airways through neutrophil DNA lysis at sites of infection. Its use has been extrapolated from CF management; however, mucus fluidity is usually conserved in PCD and there is only low-level evidence for its use, derived from case studies with short follow-up periods (109).

It is therefore not currently recommended in PCD and further studies are needed to justify the use of this high-cost medication. Approximately 5% of children under the care of the National Children's PCD management service in England were using rhDNase at the time of their 2015 annual review (6).

Most experts recommend prescribing oral antibiotics if worsening of respiratory symptoms or deterioration of lung function is observed (28). Choice of antibiotic should be directed, when available, by culture results and microbial sensitivity. While most exacerbations can be successfully treated with oral antibiotics and increased frequency of airway clearance techniques, severe or refractory exacerbations might require intravenous antibiotics and hospitalisation. Regular intravenous three-monthly regimes can be considered in patients that are not improving.

Antibiotic prophylaxis should be considered if repeated courses of antibiotics are required and its use should be balanced against the risk of potentially increasing antimicrobial resistance. The first multicentre, double-blind, parallel group RCT on pharmacotherapy in PCD showed that six-month regimen with azithromycin significantly decreased the rate of pulmonary exacerbations by almost half during the treatment period compared to placebo and was well tolerated by patients (unpublished data). In England, approximately 60% of children under PCD care in 2015 were on prophylaxis, of which Azithromycin was the antibiotic prescribed most often (6).

Treatment of asymptomatic patients with positive culture remains controversial. The most frequently isolated pathogens in children are *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Moraxella catarrhalis*. Two thirds of 333 children with PCD in England cultured a respiratory pathogen at annual review, of which 54% were *Haemophilus influenzae*. However, there was no statistically significant difference in FEV_1 z-scores between those that cultured a respiratory pathogen and those that did not grow any pathogens on culture (6).

A recently developed expert consensus statement defined pulmonary exacerbations as the presence of three or more of the following items: increased cough; change in sputum volume and/or colour; patient- or parent-perceived increased shortness of breath; decision to initiate or change antibiotic treatment due to pulmonary symptoms; malaise, tiredness, fatigue or lethargy; new or increased haemoptysis; and temperature above 38^{o}C (110). These were reduced from an initial list of 21 criteria and only included symptoms to avoid dependency of access to tests due to geographical dispersion.

Lung transplantation might be an option in cases of end stage lung disease, despite the lack of guidelines and no clear criteria for transplant referrals in patients with PCD.

2.5.1.2 Monitoring

Spirometry is often used to monitor lung disease progression since it is widely available and easy to perform in patients over five years of age; however, studies have questioned the sensitivity of forced expiratory volume in the first second (FEV_1) in PCD, particularly to detect early lung damage (111, 112). A recently published meta-analysis on spirometric indices in PCD included 24 studies published up to 2017, of which only seven had longitudinal measurements of lung function (113). They found that FEV_1 % predicted values ranged from 51% to 96%, with a weighted mean of 75%. When stratified by children and adults, subgroup analysis showed that the weighted mean in adults was lower than in children (63% versus 81%, respectively). However, there was a high heterogeneity between studies included in the meta-analysis, even after adjusting for available explanatory factors in a meta-regression.

A large retrospective cohort study with 991 patients from 21 PCD centres reported lower mean FEV_1 and forced vital capacity (FVC) at diagnosis in all PCD patients, regardless of age and country, compared to GLI reference of healthy individuals, adjusted for age, sex, height and ethnicity. Almost half of PCD patients presented with an FEV_1 z-score below the lower limit of normality (*i.e.* -1.64 z-scores) (114). Lung function seem to deteriorate with age, with children presenting higher values of FEV_1 z-scores compared to adults, and there was no correlation between FEV_1 and age at diagnosis; however, data were not longitudinal. Another recently published study reported a mean of -1.9 FEV_1 z-scores in 333 children under specialist care in the four National Children's PCD Management Service in England (see 3.2.2), with significantly lower lung function seen in older children (6).

Lung clearance index (LCI) derived from MBW is a radiation-free candidate for disease monitoring and seems to be particularly sensitive to detect early lung disease compared to FEV_1 (115). LCI represents the number of lung volume turnovers needed to clear the lungs from inert gas so patients with mucus or severe inflammation will take longer to washout the gas (116). However, correlations between LCI and HRCT have shown opposing results and standardised values for PCD patients are still lacking. Further studies are needed to elucidate the role of MBW in respiratory management in PCD (117, 118).

Chest radiography should be done at diagnosis and during respiratory exacerbations, with recommendations for routine radiography in stable patients varying between every year in the UK to every two to four years in North America (103). However, it is believed to be an insensitive test to detect early structural changes (102). HRCT, on the other hand, is the gold standard imaging method for evaluation of bronchiectasis and can detect a variety of abnormalities in PCD patients early in life, including peribronchial thickening, air trapping, mucous plugging, ground-glass

opacity and subsegmental atelectasis. Most PCD centres currently use HRCT only when clinically indicated, as opposed to routine monitoring of lung structure disease, to avoid excessive radiation exposure. Correlations between bronchiectasis severity scores derived from HRCT and FEV_1 have produced conflicting results (111, 117, 119). However, HRCT scores were developed for CF and might not be directly translatable for use in PCD, underlying the need for disease-specific scores. Future studies must avoid selection bias, as patients who have routine imaging tests tend to have more severe disease. For example, in the study population described by Kennedy *et al* (119), all adults and over half of the children had bronchiectasis seen by HRCT and 27% of the study population had undergone lobectomy.

2.5.1.3 Disease progression

Studies have found contradictory results between spirometry and sex, with some reporting that women had lower FEV_1 values while others found no association (113, 114, 120, 121). However, little is understood on the mechanism. Marthin *et al* (120) found a high degree of variation in FEV_1 decline, unrelated to age at diagnosis or baseline FEV_1. Another study on adults with PCD found a steeper annual decline of FEV_1 in women and in patients chronically infected with *Pseudomonas aeruginosa* (122). However, a large adult cohort study in the UK found no difference between men and women (121). These findings suggest that pulmonary function decline is heterogeneous and might be associated with different genetic mutations, gene modifiers or environmental influences. Conclusions are still based on small studies with different designs, populations, outcome measures, diagnostic methods, and definitions of pulmonary exacerbation. It is therefore imperative to apply standardised measurements where available (*e.g.* spirometry compliant with ATS/ERS standards, pulmonary exacerbation in PCD patients) and to develop clear definitions where necessary (123). The combination of physiotherapy and treatment of pulmonary exacerbations with antibiotics might lead to recovery of baseline FEV_1 and prevent progression of lung damage, but results are contradictory (124-126).

Comparisons between lung function parameters in PCD and CF are common. The international PCD (iPCD) cohort reported the mean values of FEV_1 and FVC % predicted in PCD children were similar to those seen in children with CF up to adolescence, when CF patients had worse lung function compared to PCD (114). Recent data, comparing PCD and CF in children in England in 2015, suggested that lung function was significantly worse in PCD patients up to the age of 15 years (6). Since newborn screening for CF was introduced in England in 2007, the age at diagnosis of CF patients has decreased and initiation of treatment in early infancy might result in improved lung function (127, 128).

Both European and North American consensus statements recommend at every patient visit: a) collection of sputum culture or, in patients that are not able to expectorate, oropharyngeal cough swabs for infection surveillance; and b) evaluation of lung function to monitor disease progression (102, 103, 129). A recent study found age-dependent changes in bacterial pathogens in the lower respiratory tract, with *H. influenzae* and *M. catarrhalis* more frequent in children, and *P. aeruginosa* in teenagers and adults (130).

2.5.2 Upper airway management

Literature on management of hearing loss and chronic rhinosinusitis in PCD is scarce. A systematic review on treatment of the latter in adults with PCD found only six studies, with small samples sizes and poorly defined outcomes that varied between studies (131-135).

All PCD patients should have an initial audiology assessment at diagnosis, followed by annual ENT evaluations, or more often if required (103). Children less than six years of age should have an open field audiogram, while a pure tone audiogram is recommended for older patients. Sinus HRCT scans can be used in symptomatic patients, to assess disease severity.

Management of otitis media aims to improve hearing levels, prevent speech impairment, delays to language and academic development. The most common treatment includes temporary use of hearing aids and antibiotic therapy to cover the most common pathogens such as *Streptococcus pyogenes, Staphylococcus aureus, Streptococcus pneumoniae, Haemophilus influenzae* and *Moraxella catarrhalis*. The European consensus guidelines strongly advises avoiding myringotomy with or without insertion of ventilation tubes in children. In contrast, the North American consensus recommend counselling patients on the likelihood of need for repeated ventilation tube insertions after initial placement and occurrence of post-operative otorrhea, as the evidence against its use is derived from small observational studies (47, 102, 103, 131).

Two recent retrospective studies assessed hearing outcomes in children with PCD. Wolter *et al* (136) reported that hearing levels were maintained in a third of those treated conservatively (*i.e.* antibiotics or hearing aids) and were restored in 80% of those in whom ventilation tubes were inserted, with the groups achieving similar hearing thresholds. However, a direct comparison between management strategies was not possible as patients with worse baseline hearing thresholds systemically received the surgical intervention. Pruliere-Escabesse *et al* (48) observed a decrease in acute otitis media and an increase of chronic otitis media with age. Otitis media with effusion (*i.e.* glue ear) was prevalent throughout all age groups, persisting in over 80% of PCD patients. Ventilation tubes were inserted in half of the children, of which 50% required repeated

replacements. Conductive deafness, defined as mean air conductive loss of at least 25 decibels (dB), decreased with age but was still considerably high (46%) for 12 to 17-year-olds.

There is no clear consensus on the best treatment option for chronic rhinosinusitis in PCD (102, 103), with current management based on nasal irrigations with isotonic or hypertonic saline up to twice a day. Sinus irrigation with normal saline is often used as it is safe, cheap and normally well tolerated by patients, but studies are needed to assess its effectiveness in PCD. Similar to CF, the sinuses seem to act as a bacterial reservoir in PCD (137) and therefore bacterial eradication through endoscopic sinus surgery might prove beneficial. Alanin *et al* (138) reported a significant improvement in health-related QoL Sinonasal Outcomes Test-22 (SNOT-22) score after endoscopic sinus surgery and adjuvant treatment (*i.e.* saline nasal rinsing, topical steroids and two weeks of antibiotic treatment). Limitations of the study included the un-blinded nature of the study, the adoption of non-disease specific HRQoL outcome measures and the use of multiple interventions without assessing individual contribution of each intervention towards outcome improvement. Patients with inflammatory nasal polyps or mucosal oedema might benefit from the use of local corticosteroids.

2.5.3 Cardiac management

PCD patients should undergo investigation with a cardiac echocardiogram and abdominal ultrasound (129). A recent English study reported a prevalence of 17% of congenital heart defects in 389 patients diagnosed with PCD (49). However, studies have used different definitions and classification systems to define heterotaxy (103, 139-142) despite standardisation efforts (143). While some authors group *situs inversus* and *situs ambiguous* under the umbrella of heterotaxy (142), others believe it should be restricted to *situs ambiguous* only (141). Using the latter, Kennedy *et al* (139) reported over 6% of PCD patients presented with heterotaxy. This is likely an underestimation, as abdominal imaging and echocardiograms are not done routinely in PCD patients.

Patients with heterotaxy have increased postoperative mortality and risk for respiratory complications. Heterotaxy has been associated with ciliary dysfunction in 42% of cases (144-146). The inference is that at least some of these patients had undiagnosed PCD, suggesting that they might benefit from routine pre-surgical screening for PCD and intense post-surgical management of respiratory symptoms (146). Further studies are needed to investigate the underlying causes of this apparent association.

2.5.4 Fertility

Patients diagnosed with PCD should receive age-appropriate genetic and fertility counselling at follow-up appointments. Currently, 22 genes have been shown to be associated with male infertility (147); however, these were based on small studies and the mechanism through which these mutations affect the sperm tail structure is unclear in many cases.

There are very few studies investigating fertility in PCD. Munro *et al* (50) reported azoospermia in five and oligospermia in two of twelve men, with normal motility seen in only one of the remaining patients. New studies have pointed towards blockage of the efferent ductules as the possible cause of reduced sperm (148, 149). Raidt *et al* (150) demonstrated that motor protein composition of fallopian tube cilia is similar to that of respiratory cilia by studying fallopian tube tissue from nine women who had children through spontaneous conception.

Prevalence of infertility and subfertility in PCD is still unknown. A systematic review suggested that 58% of women and 100% of men were infertile, but the studies were highly heterogeneous (40). Recently, Vanaken *et al* (53) reported on fertility rate in 85 patients with PCD through a retrospective review of questionnaires completed by PCD patients during clinic appointments, of which 36 were women and 49 were men. Approximately half of the women had children, 39% through spontaneous conception and a further 17% after assisted fertility. Spontaneous conception was lower in the male cohort (24.5%); however, an additional 15 men benefited from assisted fertility. Fertility data were not available in half of 167 patients screened in the study, corroborating previous findings that these data are not routinely collected or reported (40).

2.5.5 Potential future therapies

In CF, early results from studies investigating the use of small molecule compounds that target the underlying defect have been encouraging (151, 152). These potentially disease-modifying therapies are aimed at defective cystic fibrosis transmembrane conductance regulator (CFTR) proteins from 242 mutations in the CFTR gene. Ivacaftor, a CFTR potentiator, was approved for commercial use following successful RCTs (153). The combination of ivacaftor and the CFTR corrector lumacaftor was also recently approved for patients with homozygous F508del mutations.

On the other hand, gene therapy in CF has been disappointing. There are difficulties in selecting appropriate vectors to overcome the inhospitable airway environment in order to deliver the replacement gene and RCTs have failed to demonstrate clinical effectiveness. Best results showed stabilisation of lung function without any significant improvement (154). Artificially engineered

nucleases for DNA editing and antisense oligonucleotides for RNA editing might offer a better option for future therapy but need further investigation as most studies are in the pre-clinical stage.

The discovery and development of therapies that can potentially rectify the basic disease defect or the resulting defective proteins in CF might serve as a model for other genetic disorders including PCD. However, the challenge is considerably higher in PCD, as it is caused by mutations in over 45 genes, with many more genes yet unidentified, and it is genetically more heterogeneous than CF. Heterogeneity in disease progression indicates that genetic variations might be a possible cause. It is therefore crucial to define the different subgroups of patients that might benefit from more aggressive or earlier treatment. Focusing on these 'at risk' patients will also enable the allocation of resources, which are limited in most clinical settings. Pushing the pipeline for new therapies requires a coordinated approach from scientists, clinicians and industry. A summary of current knowledge, recent advances and remaining barriers in PCD, with suggestions for future research and implementation of findings, is shown in Table 1.

Table 1. Summary of current knowledge, recent advances and remaining barriers in PCD. Table developed by me, previously published in the European Respiratory Review (91).

	Current knowledge	**Recent advances**	**Barriers**	**Possible action points**
Access to service	Numerous visits to medical professionals before appropriate referral (11)	Development of screening through clinical predictive tools (17, 155); ERS guidelines with clear patient referral criteria (25)	Clinical predictive tools developed in specialist setting; Limited adoption of referral criteria by GPs and general paediatricians	Validation of screening tools in appropriate settings; Dissemination to non-specialists through support groups, conferences and media

	Current knowledge	**Recent advances**	**Barriers**	**Possible action points**
Early diagnosis	Majority of PCD patients present with unexplained neonatal respiratory syndrome	Studies investigating combination of neonatal symptoms to predict PCD (39)	Limited literature on neonatal symptoms in PCD	Increase involvement of neonatologists in collaborative PCD studies
Clinical manifestations	Limited knowledge on prevalence of symptoms	Meta-analysis reporting prevalence of clinical manifestations (40)	Lack of standardised definitions for symptoms; Lack of reporting of symptoms	Develop standardised definitions for PCD symptoms; Improve reporting of symptoms in future studies
Diagnostic testing	Recommendations on diagnostic testing	Classification system derived from the ERS guidelines on diagnostic testing for PCD (25)	Lack of standardisation on test reporting; No diagnostic studies on genetic testing	Develop protocols for diagnostic test reporting; Clarify role of genetics, immunofluorescence and electron microscopy tomography on diagnostic pathway

	Current knowledge	Recent advances	Barriers	Possible action points
PCD management	Mainly extrapolated from CF and non-CF bronchiectasis (102, 103)	Limited number of prospective cohorts (124, 126, 138) and RCTs (15, 106, 108)	Lack of standardised and validated outcome measures; No definition for commonly used terms	Develop clear definition of terms and outcome measures
Future therapies	New advancements in CF on molecular and gene therapy (151, 154)	Approval of two new therapies targeting the underlying defect in CF (152, 153)	Complexity of PCD genetics	Improve gene identification; Genotype-phenotype correlations

ERS: European Respiratory Society; GPs: general practitioners; PCD: primary ciliary dyskinesia; CF: cystic fibrosis.

2.6 Rationale and aims

The literature review outlined in this chapter highlighted several gaps in knowledge on delivery of care for patients with PCD. Some of these will be investigated and addressed in this thesis.

The overarching aim is to explore various aspects of service delivery for this rare genetic disease by adopting a conceptual framework that incorporates different stages of delivery of care for patients with PCD, considering the context in which the services are delivered. These include investigating diagnostics for PCD, by evaluating the performance of a diagnostic test compared to current diagnostic guidelines, and comparing diagnostic strategies in different countries; and disease management by examining the link between disease expression and underlying genetics, comparing management strategies in different settings, and the selection of appropriate clinical outcome measurements for long-term management of this chronic condition.

Based on previous knowledge derived from the PCD literature, the following research hypothesis will be tested:

1. HSVA is an accurate test to diagnose PCD, with good inter- and intra-observer reliability when performed by scientists with extensive experience on the test (Chapter 4); and
2. Being a genetic disease, the heterogeneity of PCD phenotypes can be partially explained by the underlying genotype, with mutations in specific genes leading to the expression of more or less severe disease (Chapter 5).

Additionally, the following research questions will be investigated:

3. The absence of a core set of clinical outcome measures that can be used across studies might lead to contradictory or spurious results. In Chapter 6, I examine which clinical outcome measures have been used in PCD and what information can be derived from them. (Chapter 6)
4. Contextual evidence might impact the selection of diagnostic and management strategies. In Chapter 7, I examine which models of service delivery for PCD are in use across different countries and settings. (Chapter 7)

Chapter 3 **General methods**

Studies included in this thesis used study-specific methods; these are detailed in the corresponding chapters (see 4.2, 5.3, 6.2 and 7.2). This section covers general methods that are common to most studies. It also describes the diagnostic and management pathway in England.

3.1 Data protection and research ethics

Study data were handled under the previous Data Protection Act of 1998 and the recent EU General Data Protection Regulation (GDPR). Confidential data were kept securely in password-protected spreadsheets, separate from clinical data. Data originating from multiple centres were anonymised locally; only anonymised data were transferred to the University of Southampton for analysis, adhering to local and international data transfer agreements.

The study presented in Chapter 4 was funded by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-1214 20014). Local and national research and ethical approvals were obtained and adhered to (Southampton and South West Hampshire Research Ethics 07/Q1702/109). Videos and clinical data were anonymised at source; each centre was responsible for ensuring data were anonymous and held in accordance with UK national data protection laws. Confidential information was held in separate password-protected databases in a restricted drive at source, with regular back-up procedures for safeguarding of data. Anonymised videos were stored on the University of Southampton Research Data Storage Service, constituting the study library, which has restricted access but is available to study collaborators for training and auditing purposes.

For Chapter 5, local and national research and ethical approvals were obtained and adhered to (NRES Committee South Central Hampshire A Ethics 06/Q1702/109, London Bloomsbury Research Ethics 08/H0713/82 and Ile-de-France ethics committee CPP07729). Confidential data for the Southampton cohort were kept secure in password-protected spreadsheets, separate from clinical and diagnostic data.

Local Research Ethics Committee approval was obtained and adhered to (Ethics and Research Governance Online, ERGO reference number 20177) for the study presented in Chapter 7. We collected only anonymised data from survey participants.

3.2 Research setting

Participants included in the studies where recruited from the National Health Service (NHS) PCD diagnostic and management services in England.

3.2.1 PCD diagnostic service in England

NHS England commission three PCD diagnostic centres to provide a highly specialised service since 2006. The three centres (Southampton, Leicester and the Royal Brompton Hospital in London) work collaboratively with competency-based training programmes, difficult case meetings and annual cross-centre data auditing. The services are equipped with state-of-the-art diagnostic facilities and a multidisciplinary team (MDT) of specialist clinical and laboratory staff, consisting of clinicians, scientists specialised in HSVA, electron microscopists, physiologists, laboratory technicians and allied health professionals (3, 156).

On the patient's first diagnostic appointment, the PCD team takes a directed clinical history, measures nNO levels and obtains a nasal brush biopsy from the posterior part of the inferior nasal turbinate. Scientists in the laboratory perform HSVA on the same day the nasal brushing samples are obtained, with preliminary findings available within approximately 30 minutes of sampling. Patients should be infection-free for more than four weeks prior to nasal brushing to minimise secondary ciliary defects. Epithelial cells obtained using a cytology brush are used for HVSA, TEM, immunofluorescence and air-liquid interface cell culture. Scientists observe strips of ciliated cells suspended in buffered culture medium at high magnification and resolution by light microscopy and record images of the cilia using high-speed video equipment (500 frames/sec), which can be played back at slow speed (e.g. 30 frames/sec) for analysis of CBF and CBP. The scientists record six to ten representative videos from each nasal brushing sample to be played back at MDT diagnostic meetings.

Nasal brushing samples are also processed for TEM for ultrastructural analysis, but results take up to three months as highly trained electron microscopists examine the ultrastructure of 100 to 300 cilia and score them using a standardised scoring sheet. Clinical history and all tests performed are systematically reviewed at MDT meetings in order to reach a diagnosis.

Genetic testing has become recently available through NHS England under the following circumstances: 1) to provide a genetic diagnosis for positive PCD cases, and 2) to confirm a diagnosis in a highly suspicious case where TEM is normal (*e.g.* patients with suggestive clinical history or low nNO levels).

The three centres work similarly and share protocols for sample processing and analysis; however, there are some significant differences in equipment as highlighted in Table 2.

Table 2. Equipment used at the PCD centres for high-speed video analysis, nasal nitric oxide and electron microscopy.

	University Hospital Southampton (UHS)	**Leicester Royal Infirmary (LRI)**	**Royal Brompton Hospital London (RBH)**
Specimen slide	0.5 mm coverwell imaging chamber (Sigma-Aldrich, Poole, UK) mounted onto a glass slide	Chamber created by the separation of a coverslip and a glass slide by two adjacent coverslips	0.5 mm coverwell imaging chamber (Sigma-Aldrich, Poole, UK) mounted onto a glass slide
Microscopy	Olympus IX71 inverted microscope and condenser. Specimen inverted onto an x100 UPlan wide aperture oil objective.	Leitz Diaplan upright microscope with x100 interference contrast plan apochromat objective lens.	Leica DM-LB upright microscope with x100 oil plan objective lens
Environmental control	37^{o}C heated environmental chamber (Solent Scientific, Southampton, UK).	37^{o}C heated stage; anti-vibration table (Wentworth Laboratories Ltd, Sandy, UK).	37^{o}C heated stage; anti-vibration table (Wentworth Laboratories Ltd, Sandy, UK).
High-speed video imaging and analysis	Photron FASTCAM MC2 high-speed video digital camera and Photron software.	IDT X4 high speed camera. AVI images analysed using MotionPro software, IDT.	Troubleshooter TS-5 Fastec imaging

	University Hospital Southampton (UHS)	Leicester Royal Infirmary (LRI)	Royal Brompton Hospital London (RBH)
Nasal nitric oxide	Ecomedics CLD 88 Exhalyzer; exhalation against resistance; sampling 0.33 $L \cdot min^{-1}$	Sievers 280i Chemiluminescence; exhalation against resistance; sampling 0.3 $L \cdot min^{-1}$	Logan LR5000 Chemiluminscence Analyser (Rochester Kent); breath hold sampling 0.25 $L \cdot min^{-1}$
Electron microscope	60,000x magnification (minimum) by Hitachi H7000	60,000x magnification by Joel 1200	60,000x magnification (minimum) by Hitachi H7000

3.2.2 PCD management service in England

NHS England has commissioned four centralised National Children's PCD Management Services (Southampton, Leeds/Bradford, Leicester and the Royal Brompton Hospital in London) since 2012 and four adult PCD management centres, mirroring the paediatric services, since 2020. The service comprises of paediatric (or adult) pulmonologists, respiratory nurses specialised in PCD care, physiologists specialised in lung function assessment and respiratory physiotherapists, with access to other specialists such as ENT specialists, nutritionists, and cardiologists. Patients undergo an annual review process, where the MTD examines symptoms and test results performed throughout the previous year and draws up a personalised management plan for each patient. Most patients are followed-up at the PCD centre; however, the MDT team travels to outreach clinics for patients under shared care with local specialist respiratory teams.

Follow-up visits are held every three months (six months for adults), either at one of the four PCD centres or at local hospitals through shared care with local clinicians. Appointments include anthropometric measures (i.e. height, weight and body mass index (BMI)), spirometry, and sputum or cough swab cultures. The transition of adolescent patients to the adult service is led by the specialist nursing team, through the 'Ready Steady Go' programme (157). The programme aims to empower patients gradually, from an early age, so they can best manage their own care. A comprehensive questionnaire can serve as a checklist or reminder for healthcare professionals to cover the various domains (*i.e.* physical, psychological and social) that can be affected by the disease.

3.3 Data collection, management and analyses

Data collection was performed using standardised collection forms, as detailed in each chapter. Raw datasets for all studies included in this thesis were cleaned, coded and re-coded prior to analysis. Outliers were checked for data entry errors by cross-referencing with the original data source, where possible.

Some variables had to be re-coded due to differences in coding of raw data. For example, self-reported ethnicity was coded from raw data using coding systems specific to countries and/or settings. For Chapter 5, we used an adapted version of the Office for National Statistics (ONS) to re-code previously coded data extracted from NHS England electronic records and to code free text provided from our collaborators in France and the Netherlands before pooling data from the different datasets (158).

We made attempts to minimise missing data by collecting data from multiple data sources, where possible. These included the NHS electronic records (*e.g.* eDocs), the PCD Database and paper medical records held at UHS. Where collaborators were involved in data collection, we emphasised the importance of data completeness for analyses.

Data were analysed in STATA (version 14.0, StataCorp, College Station, TX). We checked all data for normality in order to perform the appropriate statistical tests for hypothesis testing (*i.e.* parametric, non-parametric or data transformations).

Chapter 4 Accuracy of high-speed video analysis to diagnose PCD

4.1 Introduction

4.1.1 High-speed video analysis as a diagnostic tool

High-speed video analysis (HSVA) is a technique where respiratory cilia are visualised ex-vivo with a light microscope and recorded using a high-speed video camera. Strips of cilia are assessed for parameters of ciliary function, including CBF, CBP and effective mucociliary clearance. HSVA is the only clinically available test that assesses ciliary function. It is used routinely in European and Australian centres but not as frequently in North America. In the UK, HSVA is a front-line test in all three PCD diagnostic centres (see 3.2).

Results from HSVA are available on the day of testing, while TEM, immunofluorescence, cell culture and genetic testing may take weeks or even months before results are available. As discussed in 2.5, early diagnosis of PCD is important to ensure appropriate and early initiation of treatment, potentially limiting disease progression. A survey on patient perspective on diagnostic testing in PCD across Europe found that 47% of the patients waited over a month to receive a diagnosis after being tested for PCD and only 13% were diagnosed within a week. Interviews with PCD patients revealed that they were not satisfied with the current length of time they had to wait to receive their test results (11).

Previous retrospective studies have suggested high sensitivity and specificity of HVSA as a diagnostic test. Jackson *et al* (19) found HSVA to be 100% sensitive and 93% specific to diagnose PCD in 641 consecutive patients based on MDT decision of all available tests and clinical data. However, scientists were not blinded to diagnostic outcome and HSVA was included in the reference. Papon *et al* (98) examined 34 patients with suspected PCD and found sensitivity of 96% and specificity of 95% compared to a combination of nNO and TEM. In addition to having a small sample size, the reference adopted in this study did not follow the most recent recommendations for PCD diagnosis (25).

The evidence underpinning the use of HSVA requires rigorous validation, with some questioning its accuracy and subjectivity (65, 159) due to a variety of factors, including a) secondary ciliary abnormalities, which can occur as result of infection, inflammation or damage during sampling (160); b) operator bias, as subtle abnormalities require careful evaluation by specialist with years

of extensive training and experience in detecting minor changes in beat pattern (65); and c) lack of standardisation for processing, analysing and reporting of CBF and CBP using HSVA (120). As an example of the later, laboratories have different environmental conditions for analysing CBP and CBF, some of which are known to significantly alter these two parameters, such a temperature. The lack of well-designed accuracy studies led to the conclusion that HSVA cannot confirm a diagnosis of PCD, and the evidence-based ERS guidelines could not recommend its use as a confirmatory test (25) (see 2.4.6).

If proven to be accurate in a well-designed powered and blinded study, specialist driven HSVA could assist clinicians in making an informed decision on whether to initiate treatment and counselling of patients on the day of the patients' appointment, reducing time to diagnosis and potentially limiting disease progression. Additional tests (*e.g.* TEM, genetic testing) would still be needed to diagnose complex cases and for phenotyping patients. At the time this study was conducted, genetic testing was not routinely available in the UK through NHS England and was only conducted in selected cases, primarily for research-purposes.

4.1.2 Aims

In this chapter, the accuracy of HSVA to diagnose PCD was investigated by addressing the following:

a) the sensitivity and specificity of the test to identify patients with PCD; and

b) the inter- and intra-observer repeatability of HSVA when performed by scientists experienced in the diagnostic test in order to address the question of operator dependency.

We tested the hypothesis that HSVA can accurately and reliably diagnose PCD when conducted by experienced scientists.

This study, entitled 'Accuracy of high-speed video analysis to diagnose primary ciliary dyskinesia', was published in Chest in 2019 (see 1.2).

4.2 Methodology

4.2.1 Patient population and diagnostic decisions in the clinical setting

Patients were referred to one of three English national PCD diagnostic centres for diagnostic testing between January 2015 and April 2017. Testing included a combination of clinical history,

nNO, HSVA, and TEM. In some cases, reanalysis following air-liquid interface cell culture, immunofluorescence staining, and genetic testing were also available. Investigations in the clinical setting are detailed in 3.2.

Diagnostic results were reviewed at MDT meetings, composed of clinicians, and scientists specialised in HSVA and in TEM. The MDT considered all clinical and diagnostic data to reach a diagnostic outcome; patients were classified as 'PCD positive', 'PCD highly likely', 'PCD highly unlikely', or 'inconclusive' based on clinical experience and in accordance to the latest ERS evidence-based diagnostic guidelines (see 2.4.6).

4.2.2 Inclusion and exclusion criteria for selection of study samples

Where multiple brushings were carried out in the same patient, only the first nasal brushing from a fresh sample obtained during the study period (i.e. from 2015 to 2017) was included in the study. Similarly, when more than one MDT report was available (e.g. as a result of repeat brushings), only the MDT report related to the study sample was considered for clinical diagnostic outcome.

Laboratory technicians with experience in HSVA assessed the quality of videos against criteria agreed a-priori, outlined in the study standardised operating procedure (SOP). Exclusion criteria were based solely on issues with video image quality. These included: a) absence of at least one video with a top view of the cilia; b) lack of focus or insufficient contrast; c) short running length, insufficient to allow observation of at least ten ciliary beats; and d) if the video was considered inadequate for diagnostic purposes.

After selection and anonymisation, the technicians transferred a set of six videos per patient (five side views and one top view) to the study coordinator using the Southampton Drop Off website (https://dropoff.soton.ac.uk/).

4.2.3 Study population and data

Patients seen at any of the PCD diagnostic centres during the study period were randomly selected for inclusion in the study. Clinical data were extracted from local clinical databases and entered into a study database that contained definitions and coding for each study variable. These included clinical symptoms (derived from the PICADAR clinical predictive tool), nNO results, TEM, additional tests (where available), and final diagnostic outcome by MDT decision (see Appendix A for outcome definitions) (17). Video images were anonymised and uploaded to a central platform.

Three scientists with over ten years of experience in HSVA, one from each English National PCD Diagnostic Service, independently viewed videos from anonymised patient samples (six videos per sample), blinded to any other data. The HSVA scientists were not aware of dates of the study period and were not involved in data extraction or uploading. They reported one study outcome per sample (i.e. 'PCD positive', 'PCD highly likely', 'PCD highly unlikely' or 'inconclusive') after analysing the set of six videos available for each sample.

Scientists analysed and scored the following variables, using a standardised scoring sheet: synchronisation, stiffness, and predominant ciliary beat pattern; ciliary beat frequency; mucociliary clearance; mucus, bacteria, and blood contamination; mucus induced dyskinesia; overall cell health; final outcome; reason for inconclusive outcome; and any relevant comment (see Appendix A for further information).

4.2.4 Selection of reference standards

There is no 'gold standard' reference for PCD diagnostics (see 2.4 for limitations of each diagnostic test); we therefore compared study outcomes to two imperfect references (134):

i. **outcomes defined using the ERS PCD diagnostic guidelines**, as detailed in 2.4 (25, 161) (Figure 3); and
ii. **the clinical MDT outcome as a composite reference**, extracted from contemporary MDT meeting reports.

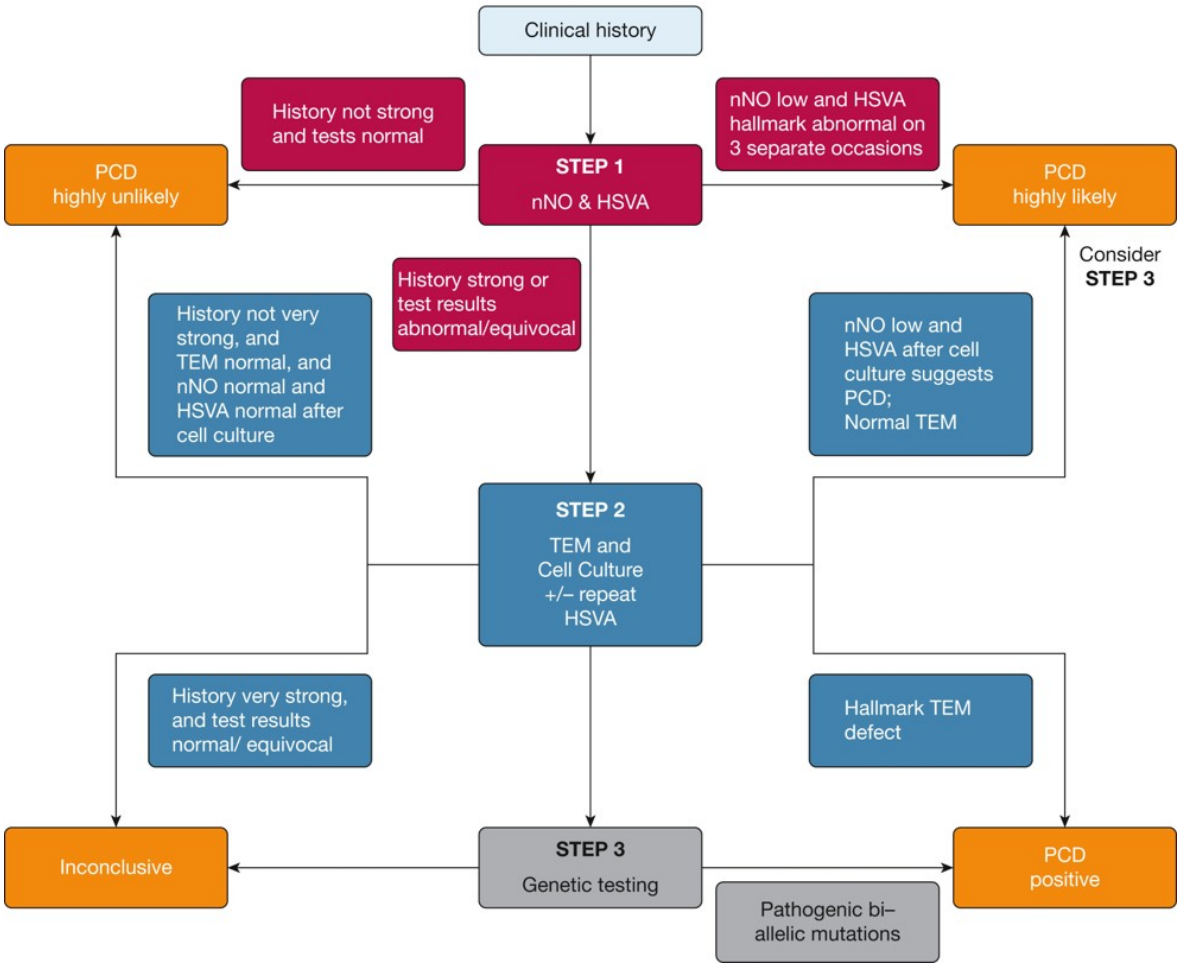


Figure 3. Schematic depicting the use of the ERS guidelines as reference standard to assess the accuracy of high-speed video analysis (HSVA). (25), from Rubbo *et al* (162).

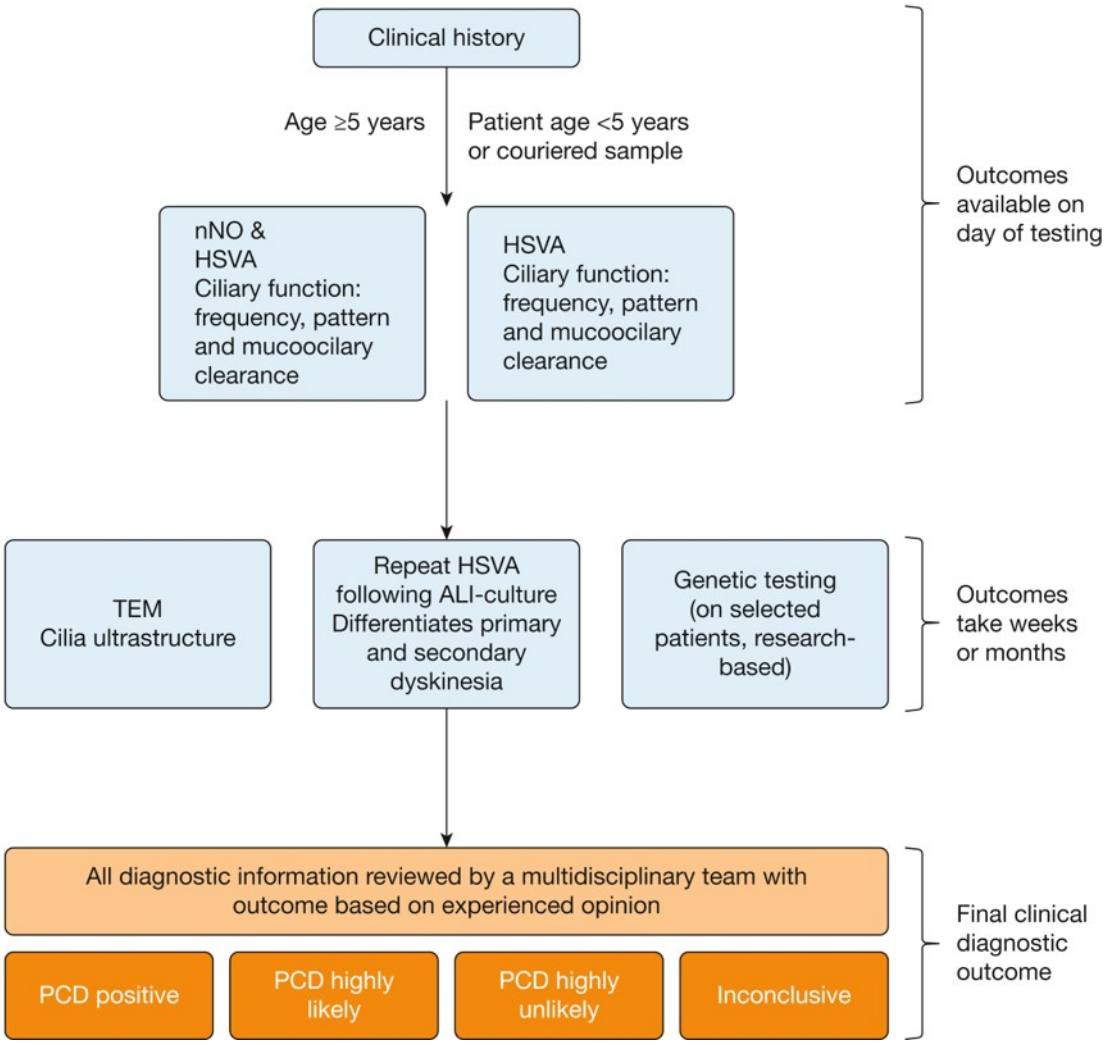


Figure 4. Schematic depicting the use of clinical multidisciplinary team outcome as reference standard to assess the accuracy of high-speed video analysis (HSVA) (26), from Rubbo *et al* (157).

For *reference* **(i)**, we used the ERS diagnostic guidelines definitions of 'PCD positive', 'PCD highly likely', 'PCD highly unlikely' or 'inconclusive'. All diagnostic tests and clinical data were examined retrospectively to determine the diagnostic outcome. Both 'PCD highly likely' and 'inconclusive' outcomes were considered as indeterminate when calculating test accuracy, as they do not provide a definitive outcome according to evidence-based guidelines (25, 161). 'PCD positive' outcomes were based solely on hallmark TEM (as discussed in 2.4) and/or pathogenic bi-allelic or X-linked mutations in PCD genes, and therefore did not include HSVA in the reference standard. However, TEM and genetic analysis have poor sensitivity (0.8 and 0.7, respectively) and will therefore 'miss' a significant proportion of true PCD patients.

For *reference* **(ii)**, diagnostic outcomes were extracted from the contemporary clinical MDT meeting reports, based on all data available to an expert MDT at the time of the diagnostic meeting. This reference therefore included all available clinical and diagnostic information,

reflecting current practices in the English PCD diagnostic services. However, HSVA was included in both comparator and reference. The limitations of each reference reflect the complexities of PCD diagnosis.

4.2.5 Statistical analyses

All patients referred to one of the three English diagnostic service during the study period were reviewed for inclusion. Patients were stratified into different groups according to their clinical diagnostic outcome, based on the MDT final report: 'PCD positive' (which included 'PCD highly likely'); 'PCD highly unlikely'; and 'inconclusive'. As the number of patients coming through the diagnostic centres that turn out to have PCD is considerably smaller than those that do not have PCD (8% of cases referred to UHS according to the literature), disproportionate sampling was used to enrich the number of PCD samples included in this study (61, 163). Therefore, 50% were selected from the PCD positive or PCD highly likely strata, 30% from the PCD highly unlikely, and 20% from the inconclusive stratum. The distribution of outcomes was decided prior to the start of the study, based on clinical knowledge of referrals to the English PCD diagnostic services.

The sample size calculation was based on the number of positive cases needed to determine test sensitivity. Taking disproportionate sampling into account, we used the estimated sample size in stratified 2x2 tables using the Cochran-Mantel-Haenszel test, setting the proportions for each strata and the difference in proportions between the two main strata, positive and negative (STATA command: power cmh 0.5 0.3 0.2, or(0.2)). 90 patients were needed to detect a sensitivity of 90% with +/- 0.9 confidence intervals. To allow for missing data and indeterminate outcomes, which were highlighted as a result of conducting the pilot study, videos from 120 patients were included in the study: 59 'PCD positive' or 'PCD highly likely', 36 'PCD highly unlikely', and 25 'inconclusive'. Randomisation for each diagnostic outcome stratum was performed in STATA.

To calculate the accuracy of HSVA, outcomes by each of the scientists were compared to the patient reference outcome, using **(i)** the ERS PCD diagnostic guidelines, and **(ii)** the original MDT report. For reference (i), true positives were defined as ['PCD positive' by scientist] and ['PCD positive' by reference]. Similarly, true negatives were defined as ['PCD highly unlikely' by scientist] and ['PCD highly unlikely' by reference]. For reference (ii), 'PCD positive' and 'PCD highly likely' outcomes were grouped, as clinically they are managed in a similar way and the 'PCD highly likely' group probably includes true PCD patients with normal TEM where the genotype has not yet been resolved. True positives were therefore defined as ['PCD positive' or 'PCD highly likely' by scientist] and ['PCD positive' or 'PCD highly likely' by MDT decision]. True negatives for reference

(ii) were defined as described for (i). False positive and false negative were determined when HSVA scientists did not agree with reference (Figure 5).

Inter-observer repeatability was calculated using Fleiss kappa (κ) coefficient for each diagnostic outcome (164). Twenty randomly selected patients were selected for re-analyses using the same proportions for each outcome stratum. Scientists re-examined the set of 6 videos per patient one year after initial study assessment in order to determine the intra-observer agreement. This was calculated for each of the scientists using Cohen κ coefficient, with bootstrapped confidence intervals (*n*=5). Bootstrapping can be used to calculate robust confidence intervals when variables have more than two levels, if three or more raters are involved, or if sample sizes are small (less than or equal to 30) (165, 166).

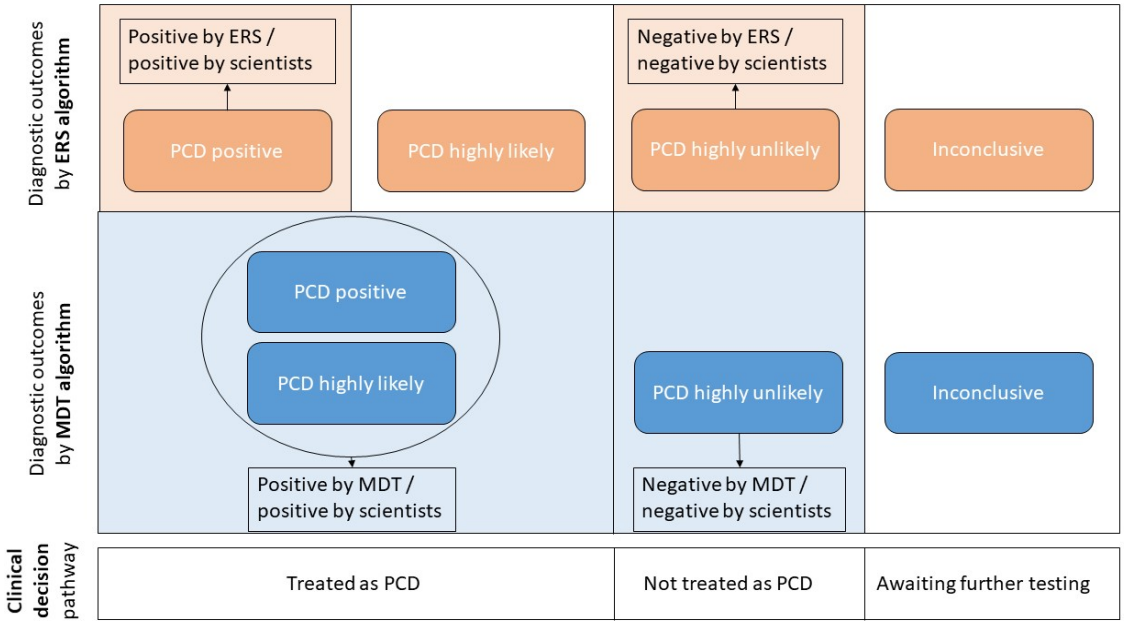


Figure 5. Schematic depicting the diagnostic algorithms used to determine 'positive' and 'negative' outcomes used to calculate sensitivity and specificity, according to the ERS guidelines (top, in pink) and the MDT final report (bottom, in blue) (25).

Continuous variables were presented as median and inter-quarter range (IQR), categorical variables were reported as proportions, and groups were compared using Fisher's exact test as the expected value for some of the outcomes was less than five. P-values were reported, with significance set at 0.05.

Sensitivity and specificity were presented with 95% confidence intervals (95% CI), where appropriate. In order to determine test accuracy, the aggregate sensitivity and specificity of HSVA outcomes against both references were calculated by combining data from each of the scientists. Data were clustered at the level of analysis (i.e. sample), as we had three outcomes for each

sample included in the study, one from each scientist. Therefore, sample outcomes are positively correlated so we modelled our data using generalised estimating equation (GEE) in order to adjust for clustering of data (167). Unadjusted models would likely lead to biased estimates of sensitivity and specificity. Importantly, this method provides robust confidence intervals for the findings, as ignoring the correlation between outcomes per sample yields misleading small confidence intervals because it counts each outcome as independent observations, inflating the sample size. Test accuracy was also calculated using worst-case scenario approach, where all 'inconclusive' outcomes by scientists were recoded as either 'false positives' or 'false negatives'. These were then adjusted for clustering using GEE modelling.

Results were reported following the STARD 2015 guidelines on reporting of diagnostic test accuracy studies (168).

4.2.6 Pilot study

A pilot study was conducted to test each stage of the project, including the standard operating procedure (SOP) for video selection, anonymisation and transfer. Nine patients from each centre were randomly selected using disproportional sampling (see 4.2.5) from referrals to the three PCD diagnostic services during 2014. Laboratory technicians retrieved videos for selected patients to evaluate video quality and excluded those that did not meet the eligibility criteria according to the SOP. Scientists received anonymised videos from 27 patients and used the standardised scoring sheet to assign one diagnostic outcome per patient.

Piloting the study highlighted difficulties in obtaining some of the clinical data, such as cardiac situs, and allowed for the implementation of a revised plan for data collection. This included requesting clinicians to evaluate radiological images to determine cardiac situs where these data were missing. It also highlighted the need to further define neonatal symptoms in order to standardise data collection of clinical history across the three centres.

As a result of the pilot study, the sample size was increased as some of the cases originally reported as positive by MDT were reclassified as inconclusive or highly likely according to the ERS diagnostic guidelines (25).

4.3 Results

The three diagnostic centres received 1285 referrals from January 2015 to April 2017. From these, 115 were considered PCD positive after review by the MDT, 852 were negative and 305 were inconclusive. Thirteen nasal brushing samples were deemed insufficient for analysis.

Characteristics of the patients whose videos were randomly selected for the study are outlined in Table 3. Genetic results were available for 16 patients, of which nine showed bi-allelic or X-linked pathogenic mutations in a PCD-causative gene (three in *DNAH5*, two in *DNAH11*, two in *CCDC40*, one in *RSPH9* and one in *OFD1*).

Table 3 Clinical characteristics of study participants, stratified by the diagnostic outcome according to the European Respiratory Society diagnostic guideline (25).

	All patients (*n*= 120**)**	**PCD positive, classified using the ERS guidelines** (*n*= 36)	**PCD highly likely, classified using the ERS guidelines** (*n*= 16)	**PCD highly unlikely, classified using the ERS guidelines** (*n*= 26)	**Inconclusive, classified using theERS guidelines** (*n*= 42)
Centre for diagnostic testing					
UHS (%)	40 (33.3%)	11 (30.6%)	3 (18.8%)	14 (53.9%)	12 (28.6%)
RBH (%)	40 (33.3%)	12 (33.3%)	3 (18.8%)	3 (11.5%)	22 (52.4%)
LRI (%)	40 (33.3%)	13 (36.1%)	10 (60.5%)	9 (34.6%)	8 (19.1%)
Median age in years (IQR)	9.6 (2.8 to 16.7)	9.1 (3.0 to 20.9)	11.8 (8.9 to 12.6)	10 (2.0 to 29.5)	7.3 (2.9 to 14.8)
Pre-term gestation (*n*=105, %)	9 (8.6%)	0	3 (23.1%)	3 (14.3%)	3 (8.1%)
Chest symptoms in neonatal period (*n*= 117, %)	97 (82.9%)	26 (78.8%)	15 (93.8%)	18 (69.2%)	38 (90.5%)
Admission to neonatal unit (*n*=109, %)	45 (41.3%)	17 (53.1%)	11 (78.6%)	7 (26.9%)	10 (27.0%)

	All patients (*n*= 120**)**	**PCD positive, classified using the ERS guidelines** (*n*= 36)	**PCD highly likely, classified using the ERS guidelines** (*n*= 16)	**PCD highly unlikely, classified using the ERS guidelines** (*n*= 26)	**Inconclusive, classified using theERS guidelines** (*n*= 42)
Presence of situs abnormalities (*n* = 118, %)	22 (18.6%)	16 (45.7%)	3 (18.8%)	0	3 (7.3%)
Cardiac abnormalities (*n*= 118, %)	5 (4.2%)	1 (2.9%)	0	3 (11.5%)	1 (2.5%)
Persistent perennial rhinitis (*n*= 118, %)	85 (72%)	28 (80.0%)	14 (93.3%)	13 (50.0%)	30 (71.4%)
Chronic ear or hearing symptoms (*n*= 116, %)	70 (60.3%)	20 (57.1%)	13 (86.7%)	13 (50.0%)	24 (60.0%)
Median nNO (nL·min^{-1}) (IQR); *n* data available	21.8 (7.2 to 105.0); *n*=72	9.8 (4.8 to 15.9); *n*=22	7.2 (3.0 to 63.6); *n*=11	189.2 (69.2 to 218.0); *n*=11	72.3 (19.9 to 117.8); *n*=28
TEM results	*n*= 120	*n*= 36	*n*= 16	*n*= 26	*n*= 42
Normal (%)	63 (52.5%)	2 (5.6%)$^{\alpha}$	7 (43.8%)	19 (73.1%)	35 (83.3%)
Isolated ODA (%)	14 (11.7%)	13 (36.1%)	1 (6.3%)$^{\neg}$	0	0
ODA+IDA (%)	14 (11.7%)	14 (38.9%)	0	0	0

	All patients (*n*= 120**)**	**PCD positive, classified using the ERS guidelines** (*n*= 36)	**PCD highly likely, classified using the ERS guidelines** (*n*= 16)	**PCD highly unlikely, classified using the ERS guidelines** (*n*= 26)	**Inconclusive, classified using theERS guidelines** (*n*= 42)
Isolated IDA (%)	4 (3.3%)	0	4 (25.0%)	0	0
IDA+MTD (%)	5 (4.2%)	5 (13.9%)	0	0	0
Central complex defect (%)	5 (4.2%)	1 (2.8%)$^{\beta}$	4 (25.0%)	0	0
Lack of cilia (%)	2 (1.7%)	0	0	0	2 (4.8%)
Inconclusive (%)	3 (2.5%)	1 (2.8%)*	0	0	2 (4.8%)
Not done (%)	10 (8.3%)	0	0	7 (26.9%)	3 (7.1%)

ERS: European Respiratory Society; PCD: primary ciliary dyskinesia; UHS: University Hospital Southampton; RBH: Royal Brompton Hospital in London; LRI: Leicester Royal Infirmary; ODA: outer dynein arm defect; IDA: inner dynein arm defect; MTD: microtubular disarrangement; $^{\alpha}$bi-allelic mutations in *DNAH11* gene; $^{\beta}$bi-allelic mutations in *RSPH9* gene; *X-linked mutation in *OFD1* gene; ¯TEM abnormality described as "thin ODA present", not a hallmark PCD defect according to the ERS Guideline (25).

4.3.1 Accuracy of HSVA compared to the ERS defined outcomes (i)

Applying the ERS guidelines as reference, 36 patients were deemed 'PCD positive', 16 were 'PCD highly likely', 26 were 'PCD highly unlikely' and 42 were 'inconclusive' (25). We found excellent sensitivity (100%) and specificity (96.2%; 95% CI, 91.7% to 100%) when comparing the study decisions of HSVA scientists with the diagnostic outcomes based on outcomes defined by the ERS guidelines (Table 4). Specificity results were adjusted for clustering as explained in 4.2.5; however,

it was not possible to adjust sensitivity as there were no 'false negatives' observed. Applying the worst-case scenario combined with GEE modelling by assuming all inconclusive outcomes by scientists were either false positive (for negative outcomes according to the ERS) or false negative (for positive outcomes according to the ERS), sensitivity remained high (93.3%; 95% CI, 92.0% to 100%) but specificity decreased to 67.9% (95% CI, 58.7% to 77.2%).

The proportion of patients classified as 'PCD highly likely' or 'inconclusive' was high for each of the scientists (39, 41 and 47 patients, Appendix A A.3 Table 17) and even higher for the reference (*i.e.* 58 patients were given an indeterminate outcome when using the ERS diagnostic guidelines as reference).

Table 4 Aggregated diagnostic study outcomes by the three scientists compared to the diagnostic outcome defined by ERS guidelines (25) (total outcomes *n*= 360 scientists' outcomes from 120 patients). Only the 'PCD positive' and 'PCD highly unlikely' outcomes (shaded in grey) contributed to the accuracy analyses.

	PCD positive by ERS guidelines (*n* patients= 36)	PCD highly unlikely by ERS guidelines (*n* patients= 26)	PCD highly likely by ERS guidelines (*n* patients= 16)	Inconclusive by ERS guidelines (*n* patients= 42)	Total (*n* patients= 120)
Study outcomes by HSVA scientists					
PCD positive	94	2	25	13	134
PCD highly unlikely	0	53	4	42	99
PCD highly likely	10	4	11	17	42
Inconclusive	4	19	8	54	85
Total *n* samples	108	78	48	126	360

Areas with grey shading were used for accuracy calculation. ERS: European Respiratory Society; HSVA: high-speed video analysis; PCD: primary ciliary dyskinesia; *'PCD positive' cases were those with a hallmark transmission electron microscopy defect and/or confirmed mutation on PCD gene.

When comparing individual scientists' HSVA diagnostic outcome with the diagnostic outcome based on the ERS guideline, we found 100% sensitivity and specificity for study decisions from both scientists A and C. Scientist B also had 100% sensitivity but lower specificity (80%) (Table 17 Appendix A A.3).

4.3.2 Accuracy of HSVA compared to MDT decision (ii)

Using the MDT diagnostic outcome as reference, 59 patients were considered 'PCD positive', 36 were 'PCD highly unlikely' and 25 had 'inconclusive' test results. There was excellent sensitivity (96.7%; 95% CI, 92.9 to 100%) and specificity (91.1%; 95% CI, 85.3 to 96.9%) of study HSVA analysis compared to the original MDT diagnostic outcome (Table 5). Sensitivity and specificity decreased to 85.3% (95% CI, 78.0 to 92.6%) and 67.6% (95% CI, 58.4 to 76.8%) respectively, when

accuracy was calculated using the worse-case scenario approach (*i.e.* when inconclusive outcomes by the scientists were recoded as false positive if negative outcome by MDT decision or false negative if positive outcome by MDT decision).

Table 5. Aggregated diagnostic study outcomes by the three scientists compared to the original diagnostic decision made by the MDT (total outcomes *n*=360 scientists' outcomes from 120 patients).

	PCD positive by expert MDT* (*n* patients= 59)	**PCD highly unlikely by expert MDT (*n* patients= 36)**	**Inconclusive by expert MDT (*n* patients= 25)**	**Total (*n* patients= 120)**
Study outcomes by HSVA scientists				
PCD positive	151	7	18	176
PCD highly unlikely	4	73	22	99
Inconclusive	22	28	35	85
Total *n* samples	177	108	75	360

Areas with grey shading were used for accuracy calculation. PCD: primary ciliary dyskinesia; MDT: multidisciplinary team; HSVA: high-speed video microscopy; *** Includes both 'PCD positive' and 'PCD highly likely' outcomes**

Sensitivity ranged from 95.9% to 100% and specificity ranged from 66.7% to 100% for individual scientist HSVA study outcomes (Appendix A A.3 Table 18). Twenty-five cases remained 'inconclusive' after review by MDT at the original diagnostic meeting. These were difficult clinical diagnostic cases that required further brushing and/or additional diagnostic testing e.g. genetic testing. The scientists reported a similar number of samples as inconclusive during the study (range 21 to 33) despite scientists relying on HSVA images alone while the MDT had the full range of clinical and diagnostic information at their disposal.

Two cases were classified as 'PCD highly likely' by both ERS guidelines and the MDT, but either 'PCD highly unlikely 'or 'inconclusive' by the HSVA scientists (Appendix A A.3 Table 17). The original clinical records indicated that one case had an isolated IDA defect on TEM (*i.e.* not a

hallmark abnormality) and five repeat brushings. Beat frequency varied between low and normal on different occasions and ciliary beat pattern was described as “almost normal” in the original HSVA from most brushing samples, some with observed mucociliary clearance. Two of the HSVA scientists classified this sample as ‘PCD highly unlikely’ and one deemed it ‘inconclusive’. The second case had normal diagnostic tests (nNO, TEM and genetics for known PCD genes) but was diagnosed as ‘PCD highly likely’ based on “semi-rotating” CBP coupled with a sibling diagnosed with PCD exhibiting similar clinical symptoms and HSVA findings to the patient in question. Two scientists classified this sample as ‘highly unlikely’, while one said it was ‘inconclusive’. Both cases are currently treated as PCD (*i.e.* under management care by the PCD teams) but require further testing before a definite diagnostic outcome can be determined, exemplifying the complexity in reaching a diagnosis in difficult cases.

4.3.3 Intra-observer reliability for each scientist

All scientists had good intra-observer agreement. We found that the Cohen κ agreement was 0.70 for scientist A (95% CI, 0.56 to 0.77), 0.66 (95% CI, 0.42 to 0.75) for scientist B, and 0.78 (95% CI, 0.61 to 0.85) for scientist C. Importantly, none of the scientists changed the outcome of their original assessment from positive to negative or vice-versa (Appendix A A.3 Table 19).

4.3.4 Inter-observer reliability for each diagnostic outcome

Using Fleiss kappa (κ) agreement (137) to compare scores between the three scientists for each diagnostic outcome, we found substantial agreement (κ=0.70) for ‘PCD positive’ and moderate agreement (κ=0.44) for ‘PCD highly unlikely’. Agreement was low for ‘PCD highly likely’ (κ=0.11) and ‘inconclusive’ (κ=0.20) (140). The overall agreement was moderate (κ=0.42; 95% CI, 0.41 to 0.44).

4.3.5 Evaluation of PCD cases with a genetic diagnosis

Scientists were able to identify the two samples with bi-allelic mutation on *DNAH11* as PCD through HSVA alone. As expected, the *DNAH11* cases had normal TEM. The scientists were also able to identify the two cases with mutation on the *CCDC40* gene using HSVA; the three *DNAH5* cases; and the single *RSPH9* case. TEM result was not available for the patient with *OFD1* X-linked mutation; however, the scientists identified this sample as PCD due to abnormal HSVA.

4.4 Discussion

4.4.1 Accuracy of HSVA to diagnose PCD

HSVA had excellent sensitivity and specificity to diagnose PCD when conducted by scientists experienced in performing the diagnostic test. However, a considerable proportion of patients had an inconclusive outcome and remained without a diagnosis after HSVA testing. These would require additional tests such as TEM, genetic testing, immunofluorescence and/or cell culture to provide a final diagnosis.

The use of two different references reflects the fact that no diagnostic test alone can diagnose all cases of PCD. Our study found sensitivity of 100% and specificity of 96% when using diagnosis based on the ERS guideline as a reference, and 96% sensitive and 91% specific when using the clinical diagnostic outcome derived from MDT reports as standard. This decrease in sensitivity and specificity when applying the MDT outcomes indicates that the accuracy of the test diminishes when more complex cases are considered, such as those deemed 'highly likely'. These were excluded from the accuracy calculations when using the ERS guidelines as reference but included in the calculations when using the MDT reports.

The publication of the ERS PCD diagnostic guidelines (25) shortly before the start of the study provided a unique opportunity to apply a standardised guideline that is currently in use internationally as the main reference to which HSVA study outcomes would be compared, improving the generalisability and comparability of study results.

Independently analysing 720 videos from 120 patients, HSVA scientists correctly identified all 'PCD positive' cases using the ERS guideline as reference. Considering that these patients have either a hallmark TEM or pathogenic mutations in PCD genes, these findings suggest that HSVA can accurately detect clear-cut PCD cases, when performed by experienced scientists. If those with an ERS-defined 'highly likely diagnosis' (*i.e.* lack of hallmark TEM or genetic confirmation but at least three HSVA abnormal results or two abnormal results plus abnormal air-liquid interface cell culture) were to be considered as true PCD cases, the detection rate would increase by 15% in the study population. This increase matches the approximate 20% PCD cases without a hallmark TEM defect reported in the literature (25, 26, 65, 84), suggesting that HSVA can pick-up cases that might have been otherwise 'missed' by TEM (25, 36, 84, 169). This hypothesis is further supported by the fact that PCD patients with an undetected *DNAH11* bi-allelic mutation exhibit abnormal HSVA (36). These would have been missed by TEM alone and might be misdiagnosed if samples are not sent for genetic analysis, immunofluorescence or three-dimensional electron tomography (36, 89).

Scientists reported two study samples as 'highly unlikely' or 'inconclusive', whereas both patients were classified as 'PCD highly likely' by MDT and ERS guidelines. Upon further review of the diagnostic history of these cases, clinical decisions were based on extensive repeat testing coupled with strong clinical and family histories, highlighting the complexity of these cases. Experts agree that some subtle beat pattern abnormalities are difficult to spot using HSVA, even with extensive training and years of experience (82). Additionally, secondary abnormities are common even in samples from healthy individuals, emphasising the need for experienced personnel to focus on the overall findings when analysing the whole cilia (82, 87, 97, 170). It is therefore not surprising that in our study population a high proportion of patients had indeterminate outcomes according to both ERS guidelines (35%) and MDT decisions (21%). This was also reflected in the number of 'inconclusive' outcomes by the scientists (23%).

The fact that sensitivity remained high even when reclassifying all 'inconclusive' outcomes by HSVA as false negatives is encouraging. However, it should be noted that cases deemed inconclusive by the two references were excluded from accuracy calculations and these might represent the more complex cases, which would require further testing before reaching even a preliminary diagnosis. Additionally, we found a decrease in specificity when cases classified as inconclusive by the scientists were reclassified as incorrect (*i.e.* worst-case scenario). This might be explained by the fact that the scientists were less confident to rule-out PCD based on HSVA alone, which was expected given the fact that they normally operate in a setting that includes knowledge of clinical data and results from other tests, such as nNO levels.

4.4.2 Reliability of HSVA to diagnose PCD

The study found a high inter-observer agreement for 'PCD positive' and moderate agreement for 'PCD highly unlikely' outcomes. This demonstrates reliability amongst experienced scientists when using HSVA to diagnose PCD. Low agreement for 'PCD highly likely' and 'inconclusive' outcomes was probably due to the interchangeability of these outcomes and likely reflects the level of confidence that scientists had when selecting an outcome, with some more confident in assigning a 'highly likely' outcome while others would have requested additional tests and therefore deemed the sample inconclusive. In practice, samples labelled as 'highly likely' or 'inconclusive' would both require a repeat brushing from the patient and further testing.

We also found good intra-observer agreement of samples by each scientist a year after the original study outcome assessment. The fact that the scientists were able to discriminate between positive and negative outcomes, agree on these between each other and with their own initial assessment, is key as these two extreme outcomes entail different clinical approaches.

4.4.3 Implications to diagnostics and clinical practice

Following current guidelines, nasal brushings are taken from every patient referred to a PCD diagnostic centre with a strong suspicion of PCD (*i.e.* suggestive clinical history and nNO level). Scientists can evaluate ciliary beat frequency and pattern using HSVA from nasal brushing samples on the same day of appointment, whereas processing and analyses for TEM can take weeks. This study demonstrated that experienced scientists can reliably use HVSA to diagnose some PCD patients on the day of testing.

The findings from this study provide the necessary evidence for clinicians to counsel patients and initiate lifelong treatment on the day of the patients' appointment, with the proviso that the final diagnostic outcome might change once all test results are available. However, TEM, immunofluorescence and genetic analysis will still be needed to confirm the diagnosis (25) and for deeper phenotyping (28, 38, 121). Importantly, the diagnosis remained inconclusive for a high number of patients following isolated HSVA, and these would need to wait for further diagnostic results. The study also showed that many patients have an indeterminate outcome using MDT or ERS outcomes even following comprehensive testing.

4.4.4 Strengths and limitations

This was the first powered and blinded study to assess the accuracy and reliability of HSVA to diagnose PCD when performed by experienced scientists. Previous literature has called for standardised methodology and reporting of diagnostic testing in PCD, in particular for HSVA (19, 25, 80, 82, 97). In this study, three experienced HSVA scientists independently assigned diagnostic outcomes prospectively using a standardised proforma by reviewing archived video images. Diagnostic outcomes and ciliary beat patterns definitions were agreed a-priori and applied in a standardised manner by the scientists when independently scoring the video images.

However, the study had limitations. There is no 'gold standard' reference to diagnose PCD; therefore, despite the use of combination testing as reference, we might have missed cases that are difficult to diagnose, likely classified as 'inconclusive' by both MDT and ERS guidelines. HSVA was present in both comparator and reference when using clinical outcomes from MDT reports as reference. However, in our comparison of HSVA with a positive diagnosis according to ERS guidelines, we excluded HSVA from the reference for sensitivity analyses as only 'hallmark' TEM and/or pathogenic mutations define a positive diagnosis. HSVA was still used to rule-out PCD in both references. We had limited genetic information available for the samples included in the study, which might have confirmed some of 'highly likely' or 'inconclusive' cases as PCD. Although

there is good standardisation of methods and reporting in the UK, protocols differ from those used in many centres.

Both sensitivity and specificity declined when applying the worst-case scenario approach. This likely indicates that the true accuracy of the test lies between the reported sensitivity and specificity (*i.e.* when all inconclusive tests were excluded from the accuracy calculations) and those reported under the worst-case scenario, where inconclusive results by scientists were considered incorrect. This underscores the need for additional diagnostic tests to be performed in order to provide a final diagnosis.

It should also be noted that despite working in three different English PCD diagnostic centres, the scientists discuss difficult cases annually at a cross-centre meeting, have vast experience in conducting HSVA, and have trained several scientists from other countries throughout the years. It is therefore probable that the accuracy and inter-observer reliability would not be as strong if conducted in centres with less experience in HSVA diagnosis. As the scientists participate in the MDT meeting, there is also a possibility that they might have recognised a few of the more complex cases. However, these are more likely to be inconclusive samples according to either reference and therefore not used for the accuracy calculations. In addition, the three centres have a high throughput of diagnostic cases and therefore it would be unlikely that the scientists would recognise videos from years ago.

The use of kappa statistics also has limitations. Both Cohen's and Fleiss kappa performs poorly when the marginal classification probabilities are either very small or very large, underestimating the strength of agreement (171, 172).This holds true in this dataset, as the probabilities of obtaining true positive and true negative results are considerably higher than of false positives and false negatives. Additionally, it relies on a convention for what should be considered substantial, moderate and low agreements. Like current trends with the use of p-values to determine statistical significance in hypothesis testing, the relevance of scientific findings should not rely solely on arbitrary conventions. Therefore, the actual numbers presented in Table 19 are more informative.

4.5 Conclusions

In conclusion, HSVA has an excellent sensitivity and specificity to diagnose PCD. There was also high agreement between scientists on 'PCD positive' and 'PCD highly unlikely' outcomes, and high intra-observer agreement, confirming that HSVA is a reliable diagnostic test when performed by experienced scientists.

As a result of this study, clinicians can make well-informed decisions based on HSVA results on whether to initiate treatment and counsel patients on the day of the patients' clinic appointment, whilst awaiting confirmatory results from other diagnostic tests. This would reduce time-to-diagnosis and potentially limit disease progression. There is now a need for international standardisation of sample processing, analyses and reporting of HSVA results.

Chapter 5 Genotype-phenotype associations in PCD using topological data analysis

5.1 Introduction

PCD is a genetically heterogeneous disorder as different mutations can produce similar phenotypes. While mutations in a single gene are responsible for all CF phenotypes, there are more than 45 genes known to be disease-causative in PCD (Table 20) and these only account for approximately 70% of all PCD cases (26).

In CF, genotype-phenotype correlation studies have identified groups with varying degrees of severity (173-175). There have been very few studies to date investigating genotype-phenotype associations in PCD (Table 20). Most of the data on phenotypic expressions of specific genotypes are derived from publications focused on gene discovery. These often have limited data on clinical characteristics in a small number of patients (range 5 to 16 patients) (8, 35, 78, 176, 177).

Blanchon *et al* (178) recently investigated genotype-HSVA phenotype correlations in a prospective study in 75 consecutive patients with PCD. They found that CBF was significantly higher in patients with *CCDC39* mutations compared to those with *CCDC40*. *RSPH1* mutations presented with less abnormal beating patterns and less ciliary rotation compared to *RSPH9* and *DNAJB13*. Patients with *DNAH11* mutations had an increased mean CBF. Severe biallelic mutation, defined as frameshift, nonsense or severe splice, in the *DNAH11* gene also resulted in reduced percent of cilia beating and impaired weighted distance travelled per second. Importantly, they found that milder mutations (i.e. missense and mild splice) were correlated with less abnormal HSVA parameters, independent of gene mutated.

DNAH5 and *DNAH11* are the most frequently mutated genes. Patients with mutations on either gene are described as having a 'typical' clinical phenotype, with randomisation of left-right pattern and high prevalence of clinical symptoms associated with the disease (i.e. NRDS, chronic otitis media, chronic rhinosinusitis) (36, 37, 179). Therefore, patients with *DNAH5* and *DNAH11* mutations have been used as a reference to which phenotypic findings from other genetic mutations are compared (28, 180).

Less severe phenotypes have been described in patients with *RSPH1*, *DNAH9*, and *CCDC103* mutations in single gene studies. Sixteen patients from ten families with mutations in *RSPH1* had better lung function, lower rates of NRDS and later onset of chronic wet cough compared to

patients with mutations in other genes (78). In another study, two out of five patients with *DNAH9* mutations presented with no respiratory symptoms and another two only exhibited mild respiratory phenotype, described as frequent cold-like infections. Data on respiratory phenotype was unavailable for the other patient with *DNAH9* mutation (181). Shoemark *et al* (35) found a lower prevalence of bronchiectasis and higher rate of otitis media in sixteen patients with a specific mutation (p.His154Pro) on the *CCDC103* gene.

On the other hand, patients with *MCIDAS* and *CCNO* mutations have been associated with a severe phenotype, including lung function deterioration and bronchiectasis at early age, followed by oxygen dependency and need for lung transplantation (176, 177). These genes are involved in the ciliogenesis process, as multicilin regulates *CCNO* expression. The absence of multicilin and cyclin-O leads to lack (or severe reduction) of motile cilia, observed on TEM and HSVA.

A retrospective study investigating fertility in PCD showed that patients with mutations in *CCDC39*, *CCDC40*, *DNAAF1* and *LRRC6* were more likely to be infertile, whilst those with *RSPH4A* mutations were fertile (53) (see 2.5.4 for further details). However, results were based on few patients with mutations in each gene. Normal fertility rates have been reported in patients with *CCDC114* mutations since CCDC114 protein expression is considerably lower in the testis compared to respiratory epithelial cells (8).

Patients with mutations in genes that code for the central complex apparatus components do not exhibit laterality defects as the rotary motion of the cilia in the embryonic node is not affected, whereas mutations on genes encoding dynein arm components affect left-right organ symmetry (90, 182). Best *et al* (49) showed that patients with *RSPH1*, *RSPH4A*, *RSPH9*, *CCDC65*, *CCDC164*, *HYDIN*, *MCIDAS*, *CCNO* and *RPGR* mutations had *situs solitus*, while mutations in the remaining 18 genes led to situs abnormalities.

Pruliere-Escabasse *et al* (48) found that patients (*n*=11) with central complex defect had a significantly higher rate of chronic otitis media (with and without otorrhea), acute otitis media, ventilation tube insertions, and recurrent otorrhea compared to 46 patients with dynein arm defects. However, the authors did not have access to genetic data and therefore could not draw further conclusions on genotype-phenotype associations.

Identifying patients that may require more aggressive or earlier treatment due to underlying genetics, will allow for better and targeted care, and patient counselling on issues such as likelihood of infertility. There have been no large multicentre studies investigating genotype-phenotype associations across multiple clinical parameters. Reasons include the complexity of clinical, diagnostic and genetic data, and statistical limitations such as multiple comparisons and

small sample sizes. Topological data analysis (TDA) is a hypothesis-free data-driven approach to data analysis that allows for the exploration of underlying patterns in complex datasets. It provides a new approach to investigate genotype-phenotype associations in patients with rare and complex diseases, without the constraints of hypothesis testing and traditional statistical analyses.

5.2 Study aims

The aim of this project was to identify patients that might have a more severe phenotype due to the underlying genetic mutations using a new and more flexible approach to data analysis. We hypothesise that some of the clinical and diagnostic heterogeneity seen across patients with PCD can be attributed to mutations in different genes.

5.3 Methodology

5.3.1 Data collection

Clinical and diagnostic data were collected from electronic and paper-based medical records at participating centres using a standardised data collection form with variable descriptors defined *a priori*. Clinical data collected included age, sex, consanguinity, ethnicity, weight and height at diagnosis, age at diagnosis, clinical symptoms at diagnosis (*i.e.* NRDS, wet cough, rhinitis, glue ear, FEV_1 and FVC), cardiac situs and CHD. Diagnostic data included nNO levels; TEM results; CBP findings by HSVA from side and top view, location of stiffness (if applicable), synchronicity of cilia beating, and CBF; and genetic data (location of genetic testing, genetic screening methods, gene mutated, mutation type and data on each of the mutated allele and consequent protein changes).

Genotyping was performed by next generation and Sanger sequencing. A geneticist specialised in PCD reviewed all genetic findings, classified allele mutations by their pathogenicity score according to the American College of Medical Genetics and Genomics guidelines (101), and issued a final genetic diagnostic decision for each patient after reviewing results from other diagnostic tests. Mutations were classified as homozygous, X-linked, compound heterozygous, single heterozygous, or novel candidate according to genetic findings. Only patients with homozygous mutations, X-linked mutations or compound heterozygous mutations (with compatible TEM results) in know PCD genes were considered confirmed cases for further analysis. A second geneticist reviewed classifications and genetic findings of all patients included in the datasets.

5.3.2 Inclusion criteria and study design

We included all patients with a conclusive positive genetic result available up to May 2019, seen at UHS and RBH in the UK, Trousseau, Cochin and Creteil hospitals in France, and Emma Children's Hospital in the Netherlands. Variants of unknown significance, single heterozygous mutations and novel candidates were classified as unsolved and therefore excluded from the study.

This was a cross-sectional study at the time of diagnosis. The initial dataset, termed discovery, consisted of patients seen at UHS and RBH up to July 2017. The validation dataset included data from patients who had genetic testing done at UHS and RBH from July 2017 to May 2019, and from France and the Netherlands up to May 2019.

5.3.3 Data classification tree

We divided PCD genes in groups according to the functional role of the ciliary protein components they encode, based on the literature (Figure 6, Appendix B Table 20). These gene groupings were used for initial inspection of the topological models as proxy for each gene mutated because they represent a small number of informative categories of the outcome of interest.

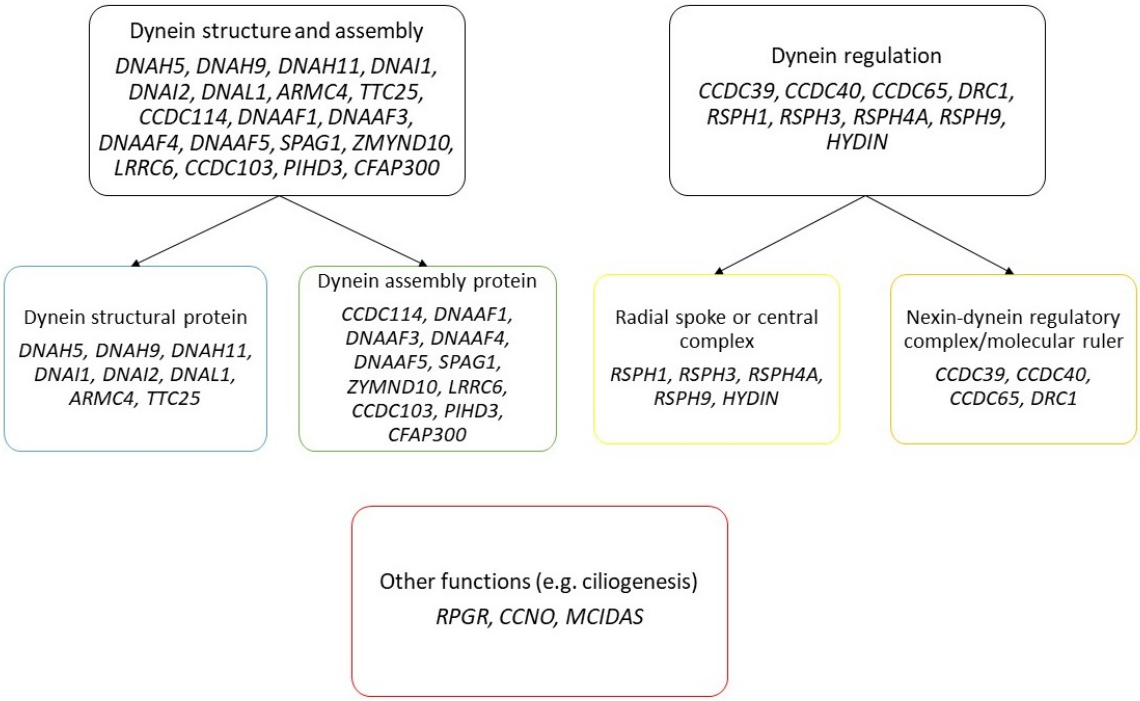


Figure 6. Schematic of stratification of PCD-causative genes into groups according to the ciliary components they encode . (*i.e.* gene groups). Each box defines a group and colours represent the gene groups: blue for genes involved in dynein structure, green for

dynein assembly, yellow for radial spoke and central complex, orange for nexin-dynein regulatory complex/molecular ruler, and red for other functions.

5.3.4 Topological data analysis

Standard cluster analysis techniques require pre-selection of parameters for clustering (*e.g.* number of clusters, variables and clustering method), which drive data analyses. In order to explore underlying patterns in complex datasets by generating clusters without the limitations of traditional clustering methods, TDA was used to generate topological models in Ayasdi.

Topology is a branch of applied mathematics that studies the shape of data. TDA consists of a set of unsupervised techniques derived from notions of similarity and distances between projected data points. It provides an unbiased and powerful approach to study the shape of data without additional model constraints (*e.g.* no assumptions of linearity or selection of feature vectors and descriptors), allowing for the exploration of the dataset in its entirety. Additionally, it can identify shape characteristics that are robust to both noise in the dataset and changes in notions of distance and (dis)similarity compared to more standard, rigid methods such as primary component analysis (PCA). The robustness of this technique is due to the fact that the outputs are not linear set of points in a low-dimensional Euclidean space but rather a combinatorial graph; therefore it can be applied to any dataset where the notion of similarity or distance is relevant to understanding the data (183, 184). Datasets derived from different times, settings and laboratories, where methodologies cannot be fully standardised, are usually difficult to interpret using traditional techniques because of noise and complexity of the data; TDA provides a good alternative for the exploration of such datasets (185).

According to topology, data have three core principles (Figure 7). The first principal is coordinate invariance, which determines that shape does not change if you alter the coordinate system used for data visualisation (*e.g.* rotate or invert the plane). The second property, known as deformation invariance, specifies the unchangeable nature of data if stretched or squashed without tear. The final property dictates that shape is maintained when using approximate geometrical shapes. Two shapes are the same if and only if one can be transformed into the other by continuous deformation. According to this concept, clusters, loops (*i.e.* continuous circular segments) and voids (*i.e.* holes) are distinct topological features (184).

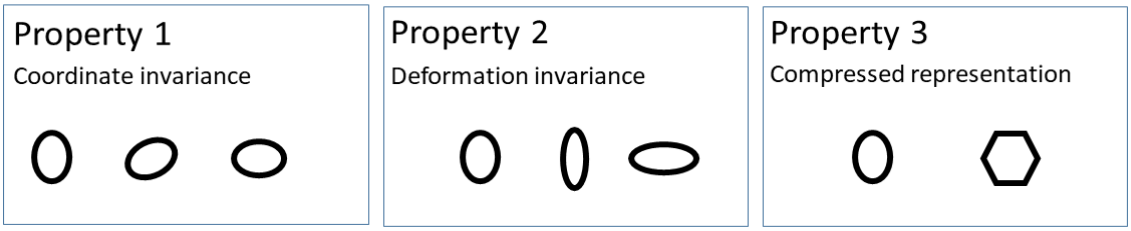


Figure 7. The three fundamental principles that govern topology. The letter "O" is still recognisable after being rotated (property 1), stretched/squashed (property 2) or compressed into a hexagon (property 3).

Ayasdi (Ayasdi Inc., Menlo Park, CA) is an online platform software that performs TDA in highly dimensional complex datasets, providing visualisation of spatial dispersion for data clustering through a hypothesis-free approach. In Ayasdi, the transformation of data points into their spatial representation (*i.e.* point clouds) is achieved with metrics that use measures of similarity (or dissimilarities) to calculate the distance between the projected points derived from the raw data and their relationship to other projected points. These are then visualised by the application of lenses or filters, which summarise the input data through a geometric, projection or statistical measure, or data feature. The application of different lenses generates networks with different shapes, allowing for the exploration of the same dataset from different mathematical perspectives. Point clouds that are close to each other, and therefore share features, are binned into nodes representing a specific cluster of data, which are then connected to other nodes where there is overlap between features. It is important to note that nodes do not represent patients but rather a subset of data points containing patients with similar features. Therefore, one node can contain multiple patients and one patient can be represented in multiple nodes (Figure 8). Two nodes are connected to each other if and only if the collection of data points contained in the nodes have a point in common (185).

Importantly, Ayasdi couples TDA with machine learning techniques, providing an even greater insight into the dataset. This is possible because machine learning techniques often generate notions of similarity, enabling the application of TDA (186), which is achieved in Ayasdi through the use of different lenses. Each lens highlights different aspects of the data shape when applied to the dataset. Ayasdi provides an automated analysis option, where statistical, mathematical or machine-learning algorithms are applied as lenses in order to produce the most suitable models in a short amount of time. These topological models can then be manually visually explored for the selection of the most appropriate one based on the outcome of interest.

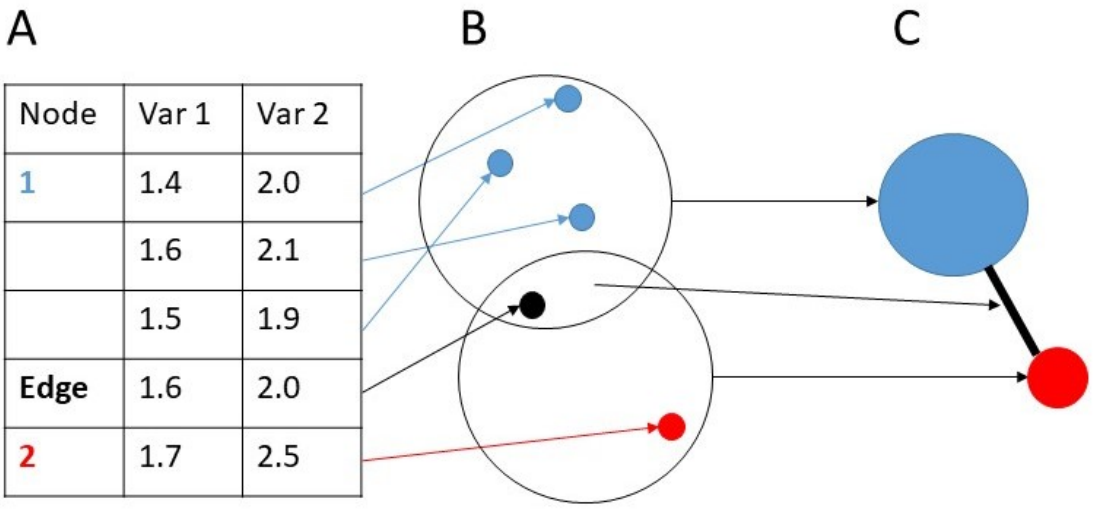


Figure 8. Schematic illustrating the generation of topological networks in Ayasdi. A) Hypothetical dataset containing five patients and two variables (Var1 and Var2); B) Projection of each data point into a point cloud; C) Transformation of the point cloud into a topological model, through clustering/binning of individual points into the blue node (four data points, three blue and one black) and red node (two data point, one red and one black) and the overlapping data point (*i.e.* edge) connecting the two nodes.

Topological models provide an interface for high-complex data visualisation and exploration. Numeric values are translated into colours, which are then applied to the nodes in the topological model. This can be continuous data, different categories that have been coded, or binary data. The process is similar to what occurs when producing heat maps.

TDA can be used to identify small groups of interest in large or complex datasets, whereas these would be 'lost' in the large volume of data when applying traditional methods (185). In doing so, it can uncover new subgroups of disease processes, severity or groups that are more likely to survive or benefit from a particular therapeutic intervention (187-189).

5.3.5 Selection of topological model

Only phenotypic data were used for clustering (*i.e.* clinical, nNO, TEM and CBP from HSVA). Genetic data were not incorporated into the analysis, as these were the study's main variable of interest. Therefore, we identified clusters solely based on the geometrical disposition of the combination of phenotypic characteristics. All other variables collected were used to explore associations with different genes by colour-coding the models according to the different variables. The automated analysis option provided several models for visual inspection, with the mutated gene groups (see 5.3.3) as main predictor of phenotypic outcome.

The topological model that showed the best well-defined clusters used locally linear embedding (LLE) lenses and the correlation distance as metric (*i.e.* distance function). LLE is a non-linear dimensionality reduction technique, on which highly complex data are summarised and compressed into smaller representations of their variability (190).

Subjects with similar features were clustered together using a resolution of 30 and gain of 3.0 (equalised). Resolution and gain are similar to camera or microscope lenses as they determine image focus, with the former responsible for the number of bins and the latter the degree of overlap between bins. Adjusting these two is an iterative process and several models were tested using different resolutions and gains, each new model informing subsequent models.

The exact same parameters (*i.e.* metric, lenses, resolution and gain) used to generate the final discovery topological model were then applied to develop the validation model so that the models would be directly comparable.

5.3.6 Statistical analysis

Outliers and potential data entry errors were checked at source, particularly for patients with extreme values of FEV_1 and FVC z-scores. BMI data were transformed (*i.e.* scaled and centred) before the application of machine learning algorithms. This step was not necessary for FEV_1 and FVC data as the use of z-scores and GLI equations provides scaled, centred and standardised values (adjusted for age, sex, ethnicity and height) (191).

Continuous variables were presented as means (with standard deviation (SD)) or medians (with interquartile range (IQR)), depending on distribution. Categorical variables were reported as proportions.

Genotypes that formed distinct clusters in the original and validation topological models were selected for comparison to the rest of the study population. Feature selection was guided by exploration of the topological models to limit the number of comparisons. The derived hypotheses were tested through statistical analyses of the whole dataset (*i.e.* merged discovery and validation datasets) in order to achieve sufficient power for appropriate statistical analysis.

Where the same outcome was tested twice, p-values were adjusted using the Bonferroni correction, with values below 0.049 considered significant. Continuous data were compared using student t-tests, ANOVA and Kruskal-Wallis, depending on the distribution of the data, and categorical data were compared using chi-square or Fisher's exact tests. Tukey's test was used for pairwise comparisons following ANOVA and Dunn's test with Holm-Sidak adjustment following Kruskal-Wallis. Multivariate regression models were used to model FEV_1 z-scores, adjusting for

age at diagnosis, history of NRDS and presence of CHD. Normality of residuals was investigated using kernel density estimations, and visual inspection of histograms and residuals versus fits graph plots. Number of observations (*n*), regression coefficients (r) with 95% confidence intervals (CI) and model's goodness-of-fitness (adjusted R^2) were reported for each model.

5.4 Results

5.4.1 Discovery dataset

The initial discovery dataset included 292 patients that were genotyped at UHS or RBH. We excluded 93 patients as they were classified as unsolved by the geneticist (see 5.3.1). The discovery topological model was therefore based on 199 patients with confirmed PCD, 48 from UHS and 151 from RBH.

The discovery dataset contained 26 different mutated genes (Figure 9). As expected, the most common genes mutated were *DNAH5* (*n*=38) and *DNAH11* (*n*=26).

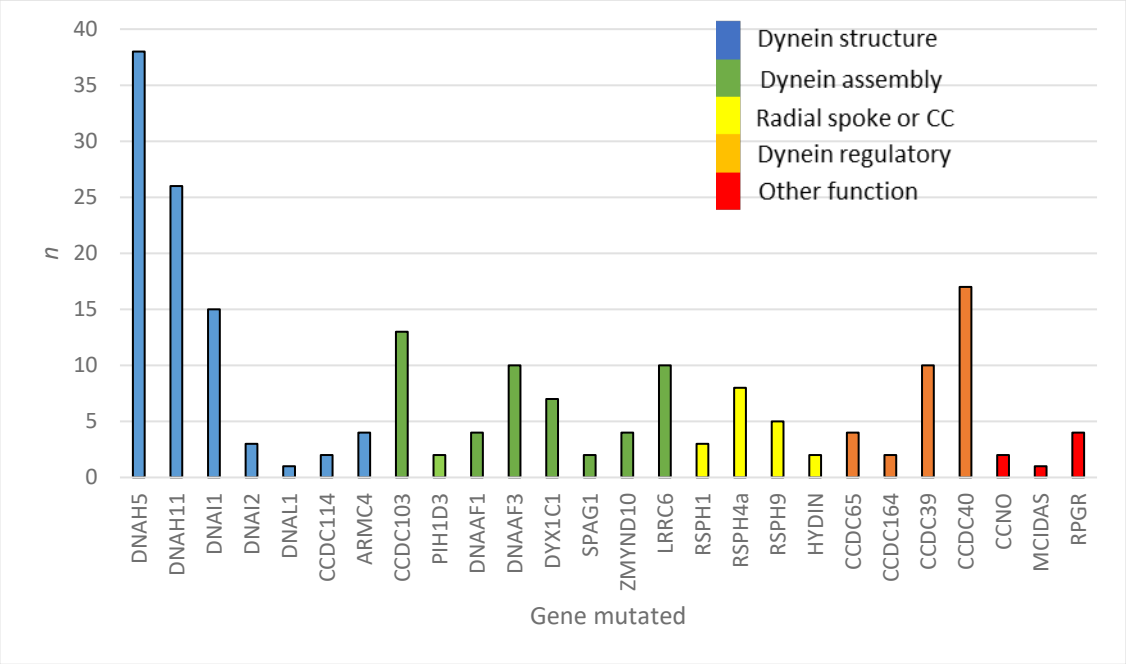


Figure 9. Distribution of genes mutated in the discovery dataset . (*n* = 199). Bars are coloured according to gene group.

5.4.1.1 Clinical and diagnostic characteristics of the discovery group

Table 6 and Table 7 show descriptive analysis on the subset of 199 patients that constituted the discovery dataset. Patients with mutation in genes that encode proteins involved in dynein structure had significantly lower nNO levels compared to those in the 'other functions' group.

TEM and CBP predominant findings were associated with the expected gene group, according to the current literature (26, 178).

Mutations in genes associated with dynein structure mostly had an ODA defect seen by TEM, while those in the 'dynein assembly' group had both ODA and IDA defects. Patients in the 'radial spoke/CC' group had central complex defects and those in the 'N-DRC/molecular ruler' group had MTD with IDA defects on TEM (Table 6). Most patients with *RPGR*, *CCNO* or *MCIDAS* mutations had lack of cilia on TEM.

All patients from the dynein structure group with a normal ultrastructure by TEM had *DNAH11* mutations (*n*=25), of whom 18 had homozygous mutations and seven had compound heterozygous mutations. The four patients with non-diagnostic TEM results in the dynein assembly group had homozygous mutations in *CCDC103* gene.

On HSVA, the predominant CBP was completely immotile for patients with genes that encode dynein structure or assembly proteins. Patients with mutations on genes that encoded radial spoke or central complex components had either a stiff (30%) or rotating (70%) beat pattern, while half of the patients with *CCDC39*, *CCDC40*, *CCDC65* or *CCDC164* had stiff cilia.

Table 6. Diagnostic characteristics of patients in the discovery group, stratified by predefined gene groups.

	Dynein structure (*n*=89)	**Dynein assembly** (*n*=52)	**Radial spoke/ central complex** (*n*=18)	**N-DRC/molecular ruler** (*n*=33)	**Other functions** (*n*=7)	**All** (*n*=199)	**p-value**
Median nNO level in nL/min (IQR); *n*=149	11.0 (6.8 to 18.8)*	17.8 (7.8 to 33.6)	23.0 (11.0 to 34.2)	12.6 (5.4 to 18.8)	39.9 (15.3 to 96.9)*	13.0 (7.4 to 24.0)	0.007
TEM findings; *n*=187							
Normal TEM (%)	25 (30.1)	4 (7.7)	4 (23.5)	2 (6.9)	2 (33.3)	37 (19.8)	

	Dynein structure (*n*=89)	**Dynein assembly** (*n*=52)	**Radial spoke/ central complex** (*n*=18)	**N-DRC/molecular ruler** (*n*=33)	**Other functions** (*n*=7)	**All** (*n*=199)	**p-value**
Isolated ODA defect (%)	51 (61.5)	9 (17.3)	0	0	0	60 (32.1)	
ODA & IDA defect (%)	6 (7.2)	34 (65.4)	0	0	0	40 (21.4)	
MTD & IDA defect or isolated IDA defect (%)	0	4 (7.7)	0	27 (93.1)	0	31 (16.6)	
CC defect (%)	0	0	13 (76.5)	0	0	13 (7.0)	
Lack of cilia (%)	1 (1.2)	1 (1.9)	0	0	4 (66.7)	6 (3.2)	
CBP predominant finding; *n*=133							
Normal (%)	0	3 (8.8)	0	0	0	3 (2.7)	
Completely immotile (%)	35 (58.3)	25 (73.5)	0	5 (22.7)	3 (42.9)	68 (51.1)	
Weak residual movement (%)	8 (13.3)	0	0	1 (4.6)	0	9 (6.8)	
Stiff (%)	16 (26.7)	6 (17.7)	3 (30.0)	11 (50.0)	2 (28.6)	38 (28.6%	

	Dynein structure (*n*=89)	**Dynein assembly** (*n*=52)	**Radial spoke/ central complex** (*n*=18)	**N-DRC/molecular ruler** (*n*=33)	**Other functions** (*n*=7)	**All** (*n*=199)	**p-value**
Rotating (%)	0	0	7 (70.0)	0	0	7 (5.3)	
Staggered beat (%)	0	0	0	5 (22.7)	0	5 (3.8)	
Lack of cilia (%)	1 (1.7)	0	0	0	2 (28.6)	3 (2.3)	

n = data available, nNO= nasal nitric oxide (normal levels <77nl/min), TEM = Transmission electron microscopy, ODA= outer dynein arm, IDA = inner dynein arm, CC = central complex, CBP= ciliary beat pattern; * significant difference between the pairs, Dunn's pairwise comparison with Holm-Sidak adjustment

Patients in the 'radial spoke/central complex' and 'other functions' groups had *situs solitus*. Those in the 'N-DRC/molecular ruler' group had significantly lower FEV_1 z-scores compared to those in the 'dynein structure' group (-2.7 vs -1.4, respectively). The proportion of patients that reported a history of NRDS (42.9%) and wet cough (71.4%) was significantly lower in those with *RPGR*, *CCNO* and *MCIDAS*; however, this group was comprised of only seven patients. History of rhinitis was reported in a significantly lower proportion of patients in the 'N-DRC/molecular ruler' and 'other functions' groups compared to those in 'dynein structure', 'dynein assembly' and 'radial spoke/CC' groups. There were no statistical differences in age at diagnosis between the five gene groups (Table 7).

Table 7. Clinical characteristics of patients in the discovery group, stratified by predefined gene groups.

	Dynein structure (*n*=89)	**Dynein assembly** (*n*=52)	**Radial spoke/ central complex** (*n*=18)	**N-DRC/molecular ruler** (*n*=33)	**Other functions** (*n*=7)	**All**	p-value
Male (%)	34 (38.2)	27 (51.9)	9 (50.0)	12 (36.4)	5 (71.4)	87 (43.7)	0.226
Ethnicity (*n*=191)							
White-British (%)	50 (58.1)	7 (13.5)	3 (16.7)	14 (50.0)	1 (14.3)	75 (39.3)	
White Irish (%)	0	5 (9.6)	3 (16.7)	1 (3.6)	4 (57.1)	13 (6.8)	
White-other (%)	10 (11.6)	4 (7.7)	1 (1.6)	5 (17.9)	1 (14.3)	21 (11.0)	
Indian (%)	4 (4.7)	5 (9.6)	0	1 (3.6)	0	10 (5.3)	
Pakistani (%)	6 (7.0)	18 (34.6)	3 (16.7)	2 (7.1)	0	29 (15.2)	
Bangladeshi (%)	0	2 (3.9)	1 (5.6)	0	0	3 (1.6)	
Sri Lankan (%)	3 (3.5)	2 (3.9)	0	0	0	5 (2.6)	
Middle Eastern (%)	1 (1.2)	1 (1.9)	5 (27.8)	1 (3.6)	0	8 (4.2)	
Black (%)	7 (8.1)	0	1 (5.6)	1 (3.6)	0	9 (4.7)	

	Dynein structure (*n*=89)	**Dynein assembly** (*n*=52)	**Radial spoke/ central complex** (*n*=18)	**N-DRC/molecular ruler** (*n*=33)	**Other functions** (*n*=7)	**All**	p-value
Chinese (%)	0	3 (5.8)	0	0	0	3 (1.6)	
Mixed (%)	1 (1.2)	1 (1.9)	0	1 (3.6)	0	3 (1.6)	
Other (%)	4 (4.7)	4 (7.9)	1 (5.6)	2 (7.1)	1 (14.3)	12 (6.3)	
Mean FEV_1 z-scores (SD), *n*=138	-1.4 (1.4) [+]	-1.9 (1.4)	-1.7 (2.1)	-2.7 (1.6)[+]	-2.7 (2.7)	-1.8 (1.6)	**0.007**
Median age at diagnosis (IQR), *n*=184	9.1 (2.0 to 23.2)	7.3 (2.3 to 12.5)	9.5 (8.4 to 15.4)	7.5 (2.0 to 13.8)	10.2 (5.8 to 12.7)	9.0 (2.9 to 15.4)	0.667
Neonatal respiratory distress syndrome, *n*=133 (%)	31 (54.4)	31 (88.6)	7 (63.6)	15 (65.2)	3 (42.9)	87 (65.4)	**0.006**
Wet cough, *n*=157 (%)	66 (94.3)	40 (100)	14 (100)	25 (96.2)	5 (71.4)	150 (95.5)	**0.042**
Rhinitis, *n*=155 (%)	65 (91.6)	38 (95.0)	11 (91.7)	18 (72.0)	5 (71.4)	137 (88.4)	**0.027**
Glue ear, *n*=144 (%)	38 (57.6)	19 (51.4)	9 (81.8)	9 (39.1)	4 (57.1)	79 (54.9)	0.206

	Dynein structure (*n*=89)	**Dynein assembly** (*n*=52)	**Radial spoke/ central complex** (*n*=18)	**N-DRC/molecular ruler** (*n*=33)	**Other functions** (*n*=7)	**All**	p-value
Situs solitus, *n*=190 (%)	31 (37.8)	19 (37.3)	18 (100)	18 (58.1)	7 (100)	93 (49.2)	**<0.001**

n = data available, [+] difference between groups was statistically significant (ANOVA followed by Tukey for pairwise comparisons)

5.4.1.2 Discovery topological model

Selection of the topological model was based on visual inspection of all generated models. We systematically explored the discovery model by colour-coding each of the variables available in the dataset. This approach was time-consuming but necessary when using a hypothesis-free technique.

Cardiac situs seemed to determine the shape of the model, as those with *situs solitus* are located on the bottom of the loop (in red), with top and bottom flares reserved to those with *situs inversus* (Figure 12 D, in blue). As there is a clear correlation between genes and determination of left-right organ pattern, this indicated that the selected topological model would be a good representation of the genes mutated.

TEM findings and CBP also mapped very closely to corresponding gene group when the model was coloured by these variables, confirming that TDA performed well according to the published literature (Figure 10, Table 20) (26). The 'dynein structure' cluster had ODA defects on TEM, and the cilia showed weak residual movement on HSVA. The genes that encode components involved in dynein assembly had combined ODA and IDA defects on TEM and were either completely immotile or had weak residual movement. Patients with genes involved in radial spoke and central complex proteins exhibited a rotating beat pattern, while those in the 'N-DRC/molecular ruler' group were stiff.

We found that patients with defects in the 'radial spoke/central complex' and 'N-DRC/molecular ruler' groups had worse FEV_1 z-scores at diagnosis (as shown in Figure 11 B in dark blue). These also mapped closely to the cluster of patients that reported a history of NRDS (as indicated in red in Figure 11 C). When investigating specific genes, patients with *CCDC39* mutations clustered in the region that exhibited lower lung function and presence of NRDS (Figure 11 D, in green). On

the other hand, those with dynein structural gene mutations had higher FEV_1 z-scores, shown in Figure 11 B in white. This cluster corresponded to those that did not report having a history of NRDS (coloured in white in Figure 11 C) and was mostly comprised of patients with mutations in *DNAH11* (Figure 11 E, in green).

When investigating upper airway symptoms, there was a distinct cluster of patients without a history of rhinitis (Figure 12 B, in white). These corresponded to the group with mutations in N-DRC/molecular ruler genes, which also have lower FEV_1 z-scores (Figure 11 B, in dark blue). The top flare of the model shows a well-defined cluster that did not report a history of glue ear (Figure 12 C in white). These were a heterogenous group composed by patients with mutations in dynein structure and assembly genes (Figure 10 A in blue and green, respectively).

Individuals with mutations in *DNAH5* were phenotypically diverse and can be seen across the entire model, with no clear cluster observed (Figure 11 F in green).

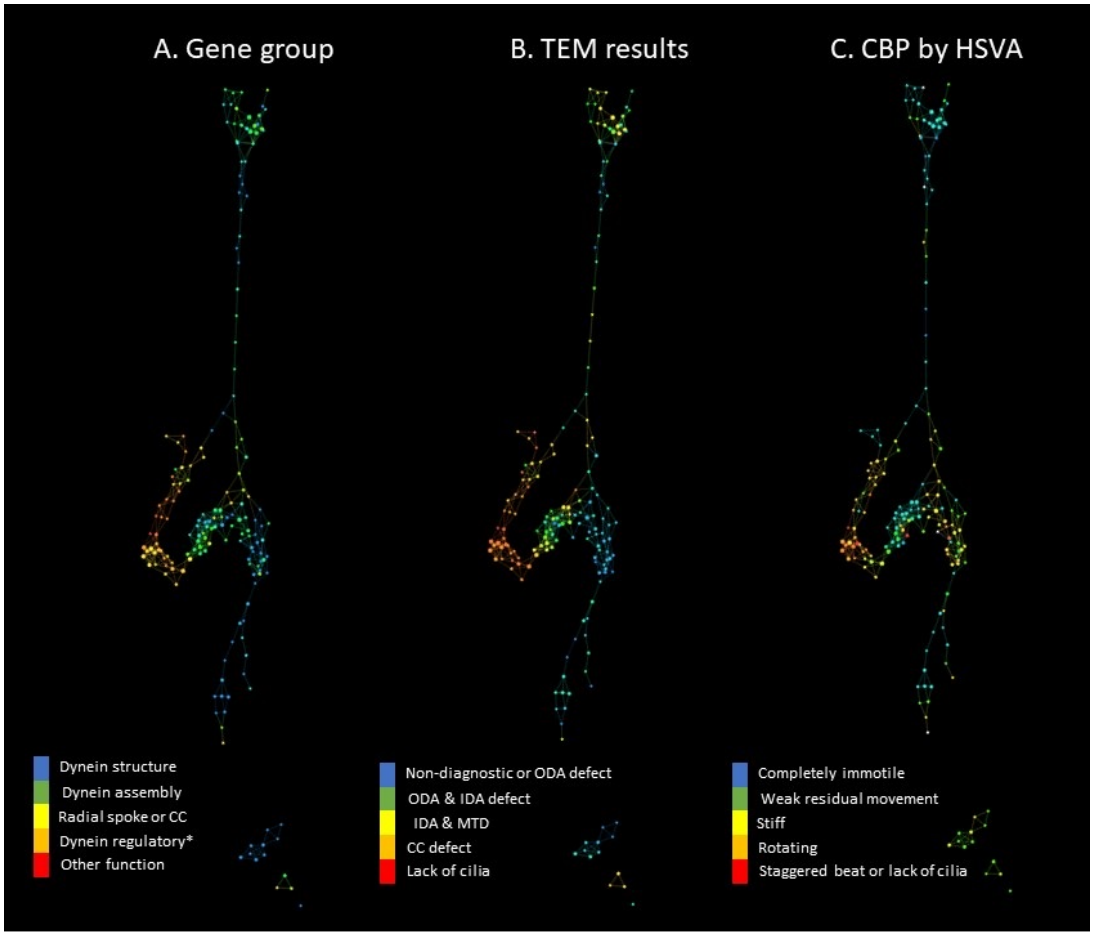


Figure 10. Topological discovery model. Models A-C are coloured by the following features: A. gene group, B. transmission electron microscopy (TEM) results, and C. ciliary beat pattern (CBP) by high-speed video analysis (HSVA).; *nexin-dynein regulatory complex/molecular ruler. Nodes represent combinations of features.

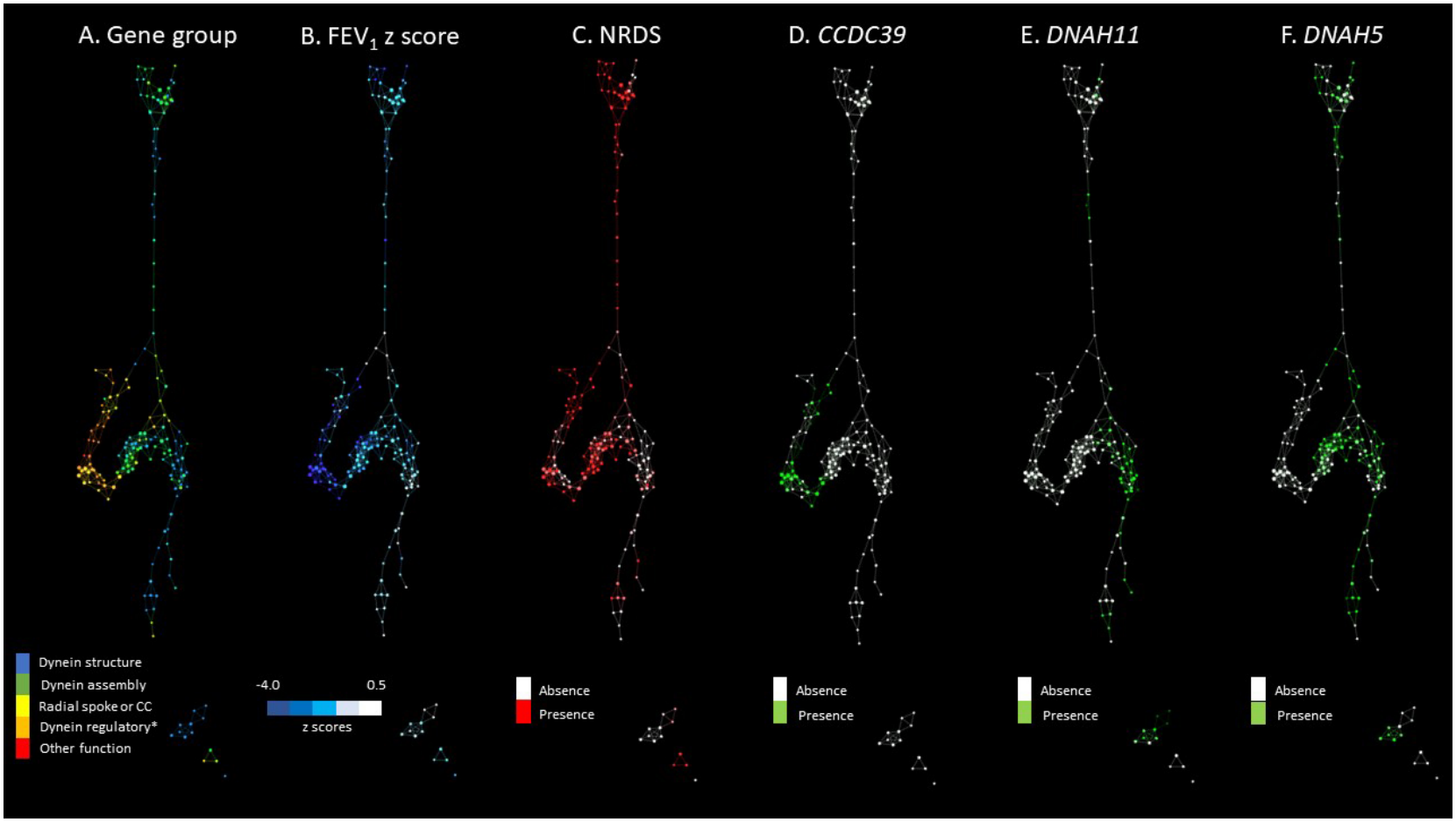


Figure 11. Topological discovery model. Models A-F are coloured by the following features: A. gene group, B. FEV_1 z-scores, C. history of neonatal respiratory distress syndrome (NRDS), D. *CCDC39* mutations, E. *DNAH11* mutations, F. *DNAH5* mutations .; *nexin-dynein regulatory complex/molecular ruler. Each node represents combinations of features.

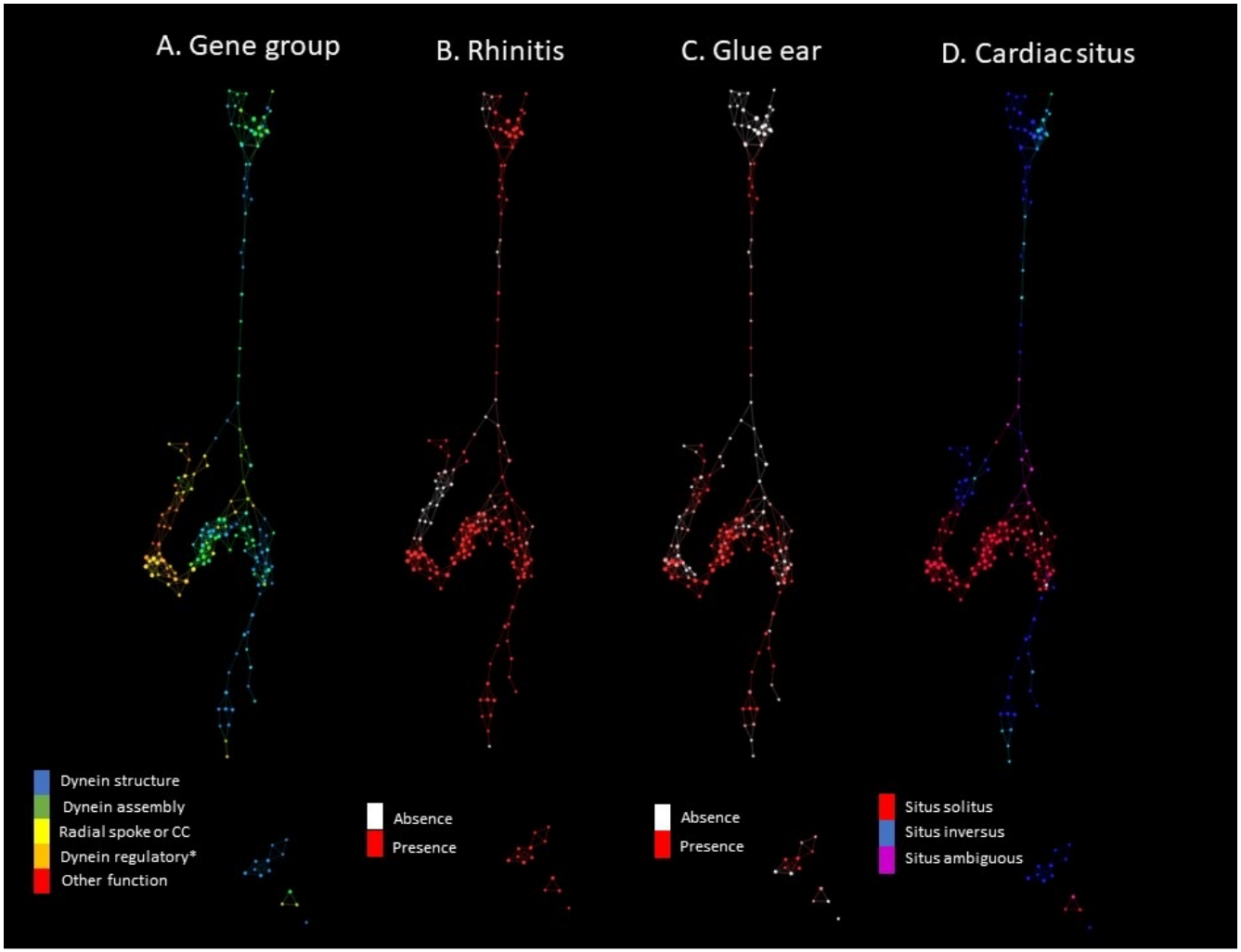


Figure 12. Topological discovery model. Models A-D are coloured by the following features: A. gene group, B. history of rhinitis, C. history of glue ear, D. cardiac situs.; *nexin-dynein regulatory complex/molecular ruler. Nodes represent combinations of features.

5.4.2 Validation dataset

The validation dataset included 108 patients from France, 61 from the UK, and 28 from the Netherlands (total *n*=197). No patients were excluded from this group as we requested that collaborators only include in the study those considered solved by a PCD-specialised geneticist (see 5.3.1 for criteria).

The validation dataset contained patients with 29 different mutated genes (Figure 13). Similar to the discovery dataset, *DNAH5* was the most frequent gene mutated, found in 40 patients; however, the second most common gene was *CCDC39*, in 25 patients.

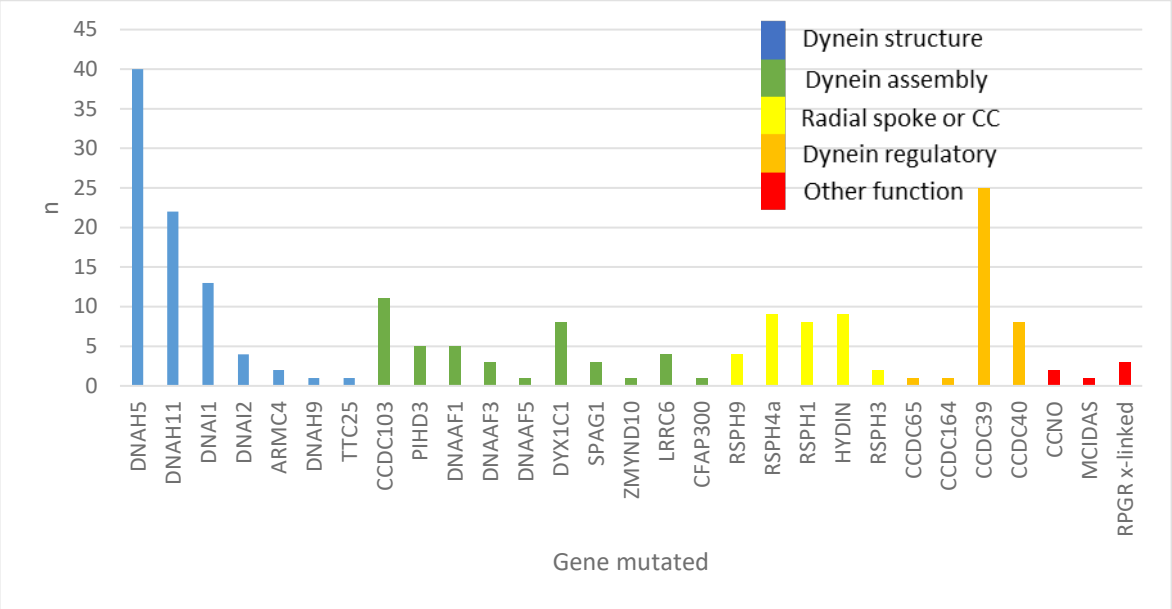


Figure 13. Distribution of genes mutated in the validation dataset . (*n* = 197). Bars are coloured according to gene group.

5.4.2.1 Clinical and diagnostic characteristics of the validation group

Findings from TEM and HSVA corresponded to the expected defects for each gene group, except for beat pattern in the 'N-DRC/molecular ruler' group. CBP in this group varied from completely immotile (42.5%), weak residual movement (36.4%) and stiff (21.2%), which might reflect differences in coding between the UK and collaborating centres in France and the Netherlands (Table 8).

Table 8. Diagnostic characteristics of patients in the validation group, stratified by predefined gene groups.

	Dynein structure (*n*=82)	**Dynein assembly** (*n*=42)	**Radial spoke/ central complex** (*n*=32)	**N-DRC/molecular ruler** (*n*=35)	**Other functions** (*n*=6)	**All** (*n*=197)	p-value
Median nNO level in nL/min (IQR); *n*=138	16 (8.1 to 23.6)	14.4 (8 to 25)	22.9 (7.6 to 40.5)	13 (9.9 to 23)	35 (15.9 to 54)	16.3 (8.4 to 28)	0.704
TEM findings, *n*=178							
Non-diagnostic TEM (%)	21 (28.4)	3 (8.3)	7 (22.6)	1 (2.9)	0	32 (18)	
Isolated ODA defect (%)	38 (51.4)	1 (2.8)	0	0	0	39 (21.9)	
ODA & IDA defect (%)	14 (18.9)	31 (86.1)	0	1 (2.9)	2 (66.7)	48 (27)	
MTD & IDA defect or isolated IDA defect (%)	0	1 (2.8)	1 (3.2)	32 (94.1)	0	34 (19.1)	
CC defect (%)	0	0	22 (71)	0	0	22 (12.4)	
Lack of cilia (%)	1 (1.4)	0	1 (3.2)	0	1 (33.3)	3 (1.7)	
CBP predominant side view, *n*=133							
Normal (%)	2 (2.6)	3 (9.1)	6 (20.7)	0	2 (40)	13 (7.4)	

	Dynein structure (*n*=82)	**Dynein assembly** (*n*=42)	**Radial spoke/ central complex** (*n*=32)	**N-DRC/molecular ruler** (*n*=35)	**Other functions** (*n*=6)	**All** (*n*=197)	p-value
Completely immotile (%)	34 (44.7)	27 (81.8)	1 (3.5)	14 (42.4)	1 (20)	77 (43.8)	
Weak residual movement (%)	29 (38.2)	3 (9.1)	6 (20.7)	12 (36.4)	0	50 (28.4)	
Stiff (%)	11 (14.5)	0	6 (20.7)	7 (21.2)	0	24 (13.6)	
Rotating (%)	0	0	10 (34.5)	0	0	10 (5.7)	
Staggered beat (%)	0	0	0	0	2 (40)	2 (1.1)	
Lack of cilia (%)	0	0	0	0	0	0	

n = data available, * nNO= nasal nitric oxide (normal levels <77nl/min), TEM = Transmission electron microscopy, ODA= outer dynein arm, IDA = inner dynein arm, CC = central complex, CBP= ciliary beat pattern.

Table 9 describes the clinical characteristics of patients included in the validation dataset. As seen in the discovery dataset, patients with mutations on genes in the 'N-DRC/molecular ruler' group had significantly lower FEV_1 z-scores compared to those in the 'dynein structure' group (-2.6 vs -1.3, respectively). However, there were no significant differences between gene groups for the other clinical parameters.

Table 9. Clinical characteristics of patients in the validation group, stratified by predefined gene groups.

	Dynein structure (*n*=82)	**Dynein assembly***)* (*n*=42)	**Radial spoke/ central complex** (*n*=32)	**N-DRC/molecular ruler** (*n*=35)	**Other functions** (*n*=6)	**All** (*n*=197)	**p-value**
Male (%)	41 (50)	22 (52.4)	14 (43.8)	23 (65.7)	4 (66.7)	104 (52.8)	0.393
Ethnicity, *n*=185							
White-British (%)	15 (20.0)	4 (11.1)	3 (9.4)	3 (9.4)	0	25 (13.5)	
White-Irish (%)	0	2 (5.6)	4 (12.5)	0	0	6 (3.2)	
White-other (%)	33 (41.8)	10 (27.8)	13 (40.6)	10 (31.3)	4 (66.7)	70 (37.8)	
Indian (%)	1 (1.3)	0	1 (3.1)	1 (3.1)	0	3 (1.6)	
Pakistani (%)	1 (1.3)	5 (13.9)	1 (3.1)	2 (6.3)	1 (16.7)	10 (5.4)	
Bangladeshi (%)	1 (1.3)	0	0	0	0	1 (0.5)	

	Dynein structure (*n*=82)	**Dynein assembly***)* (*n*=42)	**Radial spoke/ central complex** (*n*=32)	**N-DRC/molecular ruler** (*n*=35)	**Other functions** (*n*=6)	**All** (*n*=197)	**p-value**
Black (%)	2 (2.5)	3 (8.3)	1 (3.1)	1 (3.1)	0	7 (3.8)	
Chinese (%)	1 (1.3)	0	0	0	0	1 (0.5)	
Mixed (%)	5 (6.3)	0	0	1 (3.1)	0	6 (3.2)	
Other (%)	20 (25.3)	12 (33.3)	9 (28.1)	14 (43.8)	1 (16.7)	56 (30.3)	
Median FEV_1 z-scores (IQR), *n*=169	-1.3 (1.5)$^{+}$	-1.5 (1.6)	-2.1 (1.8)	-2.6 (1.5)$^{+}$	-2.6 (1.7)	-1.8 (1.6)	0.0008
Median age at diagnosis (IQR), *n*=184	14 (4.9 to 17.8)	14.3 (5.5 to 19.1)	15.9 (7.2 to 21.9)	13.9 (3.5 to 21.5)	20.4 (6.1 to 36)	14.5 (6 to 19.5)	0.435
Neonatal respiratory distress, *n*=170 (%)	41 (56.9)	21 (60)	14 (50)	20 (69)	3 (50)	99 (58.2)	0.650

	Dynein structure (*n*=82)	**Dynein assembly)** (*n*=42)	**Radial spoke/ central complex** (*n*=32)	**N-DRC/molecular ruler** (*n*=35)	**Other functions** (*n*=6)	**All** (*n*=197)	**p-value**
Wet cough, *n*=192 (%)	78 (96.3)	38 (95)	29 (93.6)	31 (91.2)	5 (83.3)	181 (94.3)	0.431
Rhinitis, *n*=192 (%)	77 (96.3)	37 (90.2)	26 (83.9)	31 (91.2)	5 (83.3)	176 (91.7)	0.150
Glue ear, *n*=187 (%)	55 (69.6)	26 (66.7)	25 (83.3)	23 (69.7)	4 (66.7)	133 (71.1)	0.574
Situs solitus, *n*=191 (%)	38 (48.1)	17 (41.5)	30 (100)	22 (62.9)	6 (100)	113 (59.2)	<0.001

n = data available, + difference between groups was statistically significant (ANOVA followed by Tukey for pairwise comparisons)

5.4.2.2 Validation topological model

The validation topological model confirmed most of the findings from the discovery model. The cluster with lower FEV_1 z-scores at diagnosis (Figure 14 B in dark blue) corresponded to patients with mutations in the *CCDC39* gene (Figure 14 D in green), whereas those with *DNAH11* mutations (Figure 14 E in green) were associated with higher FEV_1 z-scores (Figure 14 B in white). The validation model also confirmed the heterogenous distribution of patients with mutations in *DNAH5* (Figure 14 F in green).

However, there were no clear clusters with an absent history of NRDS (Figure 14 C in red). We also could not confirm the inverse association between upper and lower airway disease that was observed in the discovery model.

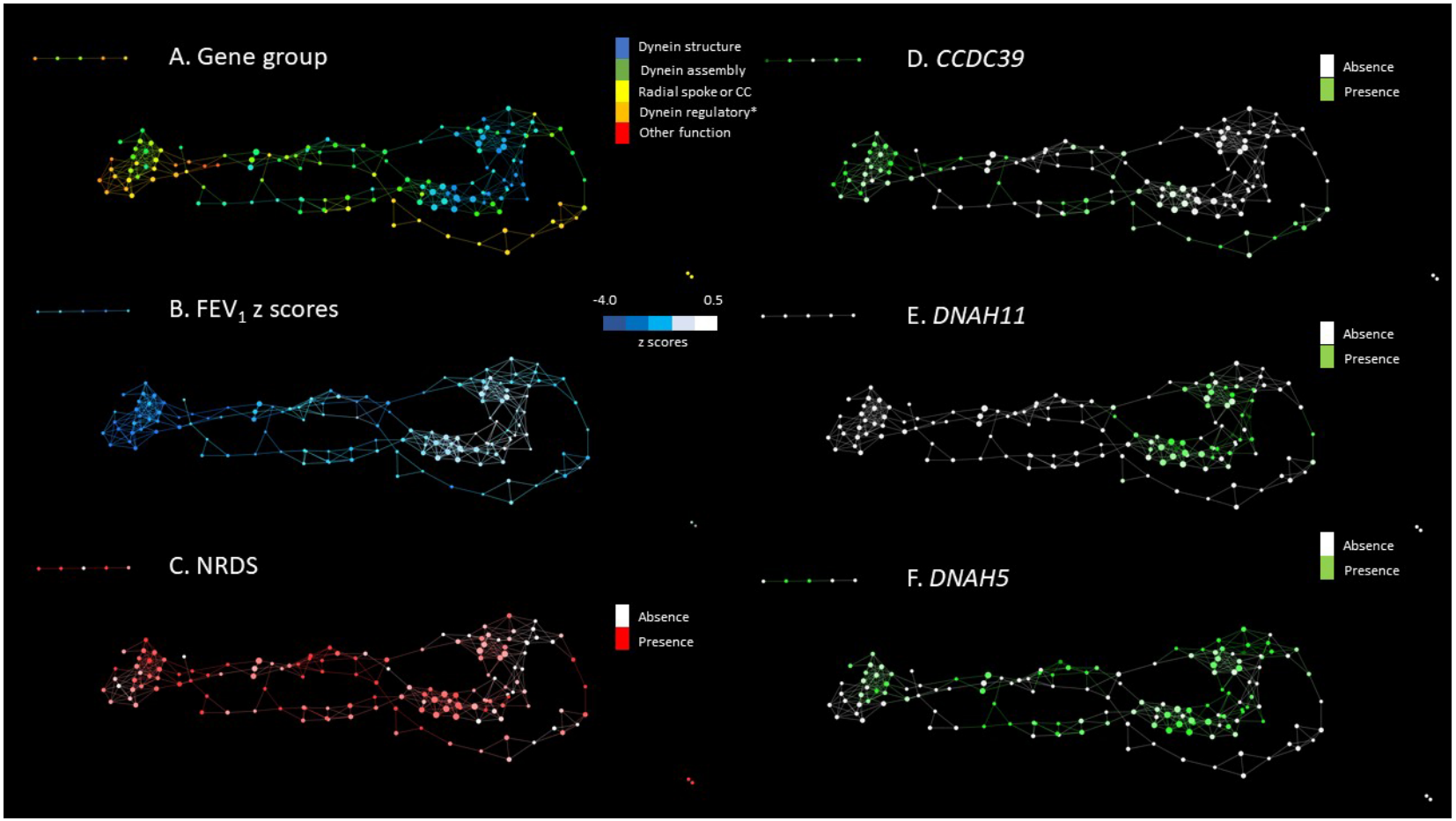


Figure 14. Topological validation model. Models A-F are coloured by the following features: A. Gene group, B. FEV_1 z-scores, C. history of neonatal respiratory distress syndrome (NRDS), D. *CCDC39* mutations, E. *DNAH11* mutations, F. *DNAH5* mutations.; *nexin-dynein regulatory complex/molecular ruler. Each node represents combinations of features.

5.4.3 Hypothesis testing of patients with *CCDC39* and *DNAH11* mutations

We selected two candidate genes for standard statistical analysis, as TDA only provides a robust method to identify groups that should be further investigated to determine what characterises them. *CCDC39* and *DNAH11* were selected due to the number of patients with mutations in each gene (*n*=35 and *n*=48, respectively) and the identification of distinct clusters in both the discovery and the validation topological models (Figure 11 D and E, Figure 14 D and E). Since these genes clustered in areas with extreme values of FEV_1 z-scores at diagnosis, we hypothesised that patients with mutations in *CCDC39* and *DNAH11* had a distinct respiratory phenotype compared to the rest of the study population.

We found that those with *CCDC39* mutations had significantly lower FEV_1 z-scores compared to patients with mutations in any other PCD gene grouped together, adjusted for age at diagnosis, history of NRDS and CHD (r = -1.2; 95%CI, -1.88 to -0.55, adjusted R^2 = 8.0%, p = 0.0004, n=205). Conversely, those with *DNAH11* mutations had significantly higher FEV_1 z-score values at diagnosis, adjusted for the same co-variables (r=0.09; 95% CI, 0.27 to 1.53; adjusted R^2=5.8%, p=0.003, n=205) , and reported less history of NRDS compared to all other patient genotypes grouped together (41.03% vs 63.91%, p=0.008).

However, we found no significant differences in NRDS in patients with *CCDC39* mutations (67.86% vs 60.29% for other genes), or in upper airway symptoms (*i.e.* rhinitis and glue ear) for patients with *DNAH11* (97.67%) or *CCDC39* (96.77%) mutations compared to any of the other genes (93.44% and 93.18%, respectively).

5.5 Discussion

This was the first large scale multi-centred European study to investigate phenotype-genotype associations in PCD across multiple diagnostic and clinical parameters. We used a novel data-driven hypothesis-free approach method to look at underlying patterns in highly complex data (*i.e.* multiple variables and outcomes) from patients with confirmed mutations in known PCD genes. We included 199 patients in the discovery dataset and 197 in the validation dataset to establish if phenotypic severity is determined by genotype.

We used TDA to locate clusters that were phenotypically homogenous and describe genes present. We then confirmed the findings from the discovery model by applying the same parameters to develop the validation topological model. Distinct phenotypic clusters where

inspected for corresponding genotype and two candidate genes were selected for standard hypothesis testing based on patterns evidenced by both models.

5.5.1 Main findings and comparison with previous literature

TDA confirmed previously reported associations between TEM, CBP by HSVA and genetics, reinforcing the role of TEM and HSVA in guiding genetic testing (25, 26, 28, 178). The diagnostic tests can also be used to confirm the likely cause of a genetic variant or to provide additional information when classifying pathogenicity of novel variants that have not yet been described or that have uncertain clinical effect.

When examining clinical phenotype, patients with dynein arm structure defects had higher FEV_1 z-scores compared to those with defects in genes that encode the radial spoke, central complex and N-DRC/molecular rulers. Those with mutations in the *CCDC39* gene had significantly lower lung function (*i.e.* FEV_1 z-scores) at diagnosis when compared to all other genes grouped together. Findings from our study confirm and add to evidence reported by Davis *et al* (28). The authors found that FEV_1 and FEF_{25-75} % predicted were significantly reduced in 24 children from six centres in the USA and Canada with IDA and MTD or with CC defects compared to 58 children with ODA or with ODA and IDA defects. They found no difference in NRDS, chronic daily wet cough, chronic nasal congestion, chronic or recurrent otitis media or laterality defects between the two groups. The majority (75%) of the patients in the IDA and MTD or CC group had biallelic mutation in *CCDC39* or *CCDC40* genes, and therefore the observed lower lung function was attributed to mutations in these two genes. No gene was identified in the remaining 10 patients in this group.

Davis *et al* (180) then followed-up these children over a period of five years in a multicentre prospective longitudinal study. They found that those with *CCDC39* or *CCDC40* mutations had a statistically significant steeper age-associated decline in FEV_1 % predicted over follow-up (-1.11% per year), lower infant FEF_{25-75} z-scores and worse growth parameters (weight and height z-scores) compared to those with *DNAH5* mutations. However, there was no significant decline in growth parameters over time.

We found that patients with mutations in *DNAH11* had significantly higher FEV_1 z-scores at diagnosis and were less likely to have had a history of NRDS compared to patients with mutations in other genes. These findings were confirmed by both the validation model and standard hypothesis testing. Irving *et al* (192) described a milder respiratory phenotype (FEV_1 z-scores and MBW-derived LCI) in 14 patients with normal ultrastructure, of which six had *DNAH11* mutations.

The discovery topological model suggested that there was an inverse association between upper and lower airway disease; however, this was not observed in the validation model. Variations in data collection and coding could account for failure to replicate results seen in the discovery model, despite the use of a standardised form, since this was a retrospective cross-sectional study. They could also reflect actual differences in population characteristics, as the discovery dataset only included patients diagnosed in the UK while the validation dataset was formed by patients from the UK, France and the Netherlands.

Patients with mutations in *DNAH5*, the commonest overall genetic cause of PCD, were phenotypically heterogenous. This is likely due to the variety of mutations within this large gene, as there are at least 100 different pathogenic variants described to date. As *DNAH5* was also the most frequent gene mutated in our datasets, it is possible that the nature (*i.e.* frameshift, nonsense, splice, missense) and location of the mutations within a gene can be as or more important in determining phenotypic severity than the actual gene mutated. This was previously suggested by Blanchon *et al* (178), as they found that mutations that allowed for the formation of proteins (*e.g.* missense and mild splice) resulted in milder ciliary beat abnormalities, while those that impeded ciliary protein formation such as frameshift, nonsense and severe splice led to more severe CBP abnormalities. Shoemark *et al* (35) showed that patients with the p.His154Pro missense mutation in the *CCDC103* gene presented with a milder diagnostic phenotype (*i.e.* normal nNO, CBF, and CBP).

5.5.2 Potential implications to clinical practice

Our study provides new insight into genotype-phenotype associations and the findings can inform education and counselling of patients with PCD. High risk groups, such as those with mutations in *CCDC39*, might benefit from closer monitoring, more regular follow-up visits and more aggressive treatment from the time of diagnosis, as they seem to have impaired lung function at an early age. Those with *DNAH11* mutations might be at a lower risk for progressive lung disease and therefore current treatments and frequency of visits could be appropriate for these patients.

5.5.3 Strengths and limitations

One of the main strengths of our study is the use of a data-driven hypothesis-free approach. TDA enabled the exploration of underlying patterns in a large multicentre dataset. Using this novel approach, we were able to identify phenotypic clusters that were distinct from each other, some of which could be traced to specific gene mutations and therefore guide the selection of genes for hypothesis testing. Furthermore, initial findings from a UK-based sample were confirmed in a

validation dataset that included patients from two other countries, suggesting that our findings can be generalised to other PCD populations.

However, a major limitation of this study is its cross-sectional nature at diagnosis; large longitudinal studies are needed to further investigate disease prognosis according to the different genotypes. Equally, the retrospective nature of our study meant that data collection was limited by what was recorded in the medical records and therefore some variables had missing data. For example, FEV_1 z-scores were missing in approximately 30% of patients, most of which were diagnosed at an early age and therefore could not reliably perform spirometry. However, TDA is particularly robust to missing data. A previous study demonstrated that even up to 90% missing data did not affect the shape of the resulting model (193).

We did not consider environmental and socioeconomic factors, differences in management strategy, and compliance with treatment, all of which likely play a significant part in disease severity. There was a potential recall bias for variables related to early life such as NRDS, which was probably more pronounced for those that were diagnosed in adulthood.

Despite including almost 400 patients with PCD, we were underpowered to tease out the phenotype of genes that were mutated in only a handful of patients; we therefore had to select two candidate genes (*CCDC39* and *DNAH11*), and combine the discovery and validation dataset to perform hypothesis testing in order to avoid multiplicity and loss of power. Larger international studies are needed to ascertain phenotypic differences for other less common genes. European collaborative networks such as BEAT-PCD (194) have identified research into genotype-phenotype correlations as a priority for future projects.

5.6 Conclusions

Using TDA and standard statistical testing, we found distinct diagnostic and clinical phenotypes that were associated to specific genes. There was a clear association between genotype groups and other diagnostic test results such as TEM and HSVA, suggesting that these can be used to inform the former when establishing a diagnosis of PCD, particularly where there are variants of unknown significance.

Patients with mutations in *CCDC39* had more severe lung impairment at diagnosis. Those with mutations in *DNAH11* had milder respiratory phenotype, with higher values of FEV_1 z-scores and less history of NRDS. Future studies are needed to explain the biological mechanism behind these findings and to explore other genotype-phenotype associations.

Chapter 6 Clinical outcome measures for use in PCD research, a scoping review

6.1 Introduction

A set of core outcome measures are the minimum outcomes that are measured and reported across all prospective studies of a specific disease or treatment. There are no agreed core sets of disease-specific outcome measures in PCD research. Yet, the choice of clinical outcome measures informs the selection of data sources from which study data should be collected; the appropriateness, frequency and length of follow-up measurements; and the required number of patients, as sample size relies on the expected frequency and variability of outcomes, and on the effect of interest (or the minimal clinically relevant difference) (146). In fact, the quality of the knowledge generated by research strongly relies on the selection of appropriate clinical outcome measures.

Outcome measures should be valid, reliable and sensitive to measuring the effect of interest (195). Validity addresses the degree to which the selected clinical outcome is accurate in measuring what it set out to measure. Sensitivity relates to the ability to detect changes in the measured outcome. Spirometry, for example, is routinely used to monitor disease progression in many common diseases but is thought to be an insensitive surrogate marker for progressive lung disease in CF and, more recently, in PCD (111, 115, 117, 196). It can be difficult to detect changes over time as decline in FEV_1 becomes relatively slow. Reliability determines the reproducibility of a measurement when repeatedly measured in the same person or sample using the same method and equipment.

In CF, an expert consensus from the CF Foundation and the US National Institute of Health (NIH) recommended four main clinical outcome measures: pulmonary function, frequency of pulmonary exacerbations, HRQoL and growth measurements (152). However, these outcomes rely heavily on the respiratory component and might not be sensitive enough to detect disease progression for patients that have no detectable sign of respiratory compromise such as younger patients, which might be precisely the population that would benefit the most from course-altering medications (116, 196). Furthermore, the magnitude of change needed to constitute a clinically significant decline is also lacking (197).

6.1.1 Disease-specific outcome measures

Well-defined and validated disease-specific outcome measures are the most efficient and accurate way to assess new therapies and management options. Whilst true for all diseases, this is particularly poignant for rare diseases, where the number of patients available is limited (189).

Disease-specific outcome measures need to be both reliable and sensitive to the selected measure. An outcome measure that is valid for another disease might not be appropriate to measure the effect of interest, or sensitive enough to detect a more subtle effect. The St George's Respiratory Questionnaire (SGRQ), for example, performs well for chronic obstructive respiratory disease (COPD) and other chronic obstructive diseases but was found to be less appropriate to measure HRQoL in PCD patients (198).

The limitations of non-PCD-specific HRQoL tools were further illustrated by results from the first crossover double blind RCT in PCD. Paff *et al* (106) found no benefit on HRQoL measures when using hypertonic saline compared to nebulised saline, assessed by the overall score from the SGRQ. However, it was unclear whether this respiratory questionnaire designed for COPD was sensitive and reliable enough to detect significant improvement in HRQoL in patients with PCD (107). This hypothesis was supported by the fact that patients reported a subjective benefit when using hypertonic saline and there was some improvement on HRQoL measured by the Quality of Life Questionnaire-Bronchiectasis (QOL-B) Health Perception Scale. Designed for patients with bronchiectasis, the QOL-B questionnaire might be appropriate for PCD patients with more advanced disease; however, it does not consider non-respiratory symptoms, such as chronic rhinosinusitis and chronic otitis media with or without hearing loss.

Chronic diseases, particularly those with an early onset, often require age-specific outcomes. Infants, children, adolescents and adults can be affected in different ways and the importance attributed to each symptom by patients can vary throughout the years, such as fertility problems in adolescents and adults, and repeated ear infections in younger children (199).

It is unclear which clinical outcome measures have been used in PCD research, and which ones are appropriate for use in prospective cohorts and RCTs.

6.1.2 Scoping reviews

Scoping reviews are similar to systematic reviews in terms of their structured systematic approach to synthesising the literature but differ in their aims (200-202). While systematic reviews focus on gathering evidence to address a specific question, the main objective of scoping reviews is to map the relevant literature in the field of interest and therefore they have a broad and descriptive

scope. Scoping reviews are also useful to a) identify key concepts and definitions in the literature, highlighting inconsistencies; b) develop specific questions that can then be addressed in systematic reviews with or without meta-analysis; and c) explore gaps in the existing literature.

After reviewing the initial data extractions, we opted to conduct a scoping review as it matched our aims and the evidence gathered by the search clearly pointed to a lack of standardisation and high heterogeneity of clinical outcome measures used, preventing any attempts to synthesise the findings through a systematic review with or without meta-analysis. We therefore focused on reporting the clinical outcome measures that have been used in PCD research and not the findings themselves, using some of the more representative studies as examples of how these outcomes have been used in the current PCD literature.

6.1.3 Study aims

The aim of this scoping review was to systematically identify, map and describe the evidence in this area, and to highlight the most commonly used outcomes, consistency of definitions across studies, and variations on the use and reporting of clinical outcome measures in the PCD literature.

6.2 Methodology

6.2.1 Search strategy

We searched Medline, Embase and Cochrane Systematic Review online databases to identify studies describing clinical outcome measures in PCD. A pilot search was conducted using only terms related to the disease (Items 1-5 of search terms, Box 1). One reviewer (BR) scanned the first 1000 abstracts to identify key terms that could be used to build the full search strategy. The search was designed for use in EMBASE with Embase Subject Headings (Emtree) and tailored for use in MEDLINE with Medical Subject Headings (MeSH). We included individual terms to both searches and applied limitations (*i.e.* human only and published after 1996). The full search strategy is described in Box 1.

6.2.2 Inclusion and exclusion criteria

Studies reporting on clinical outcome measures in PCD were included if they a) had a study population of at least 10 patients diagnosed with PCD, b) were published in English, c) were published after 1996, and d) were conducted in humans. We excluded studies prior to 1996 because the diagnostic methods and criteria have changed considerably in the last twenty years;

therefore, it was likely that the older manuscripts contained a high proportion of patients that would no longer be diagnosed with PCD following recent diagnostic consensus statements and guidelines (25, 65). However, similar to the strategy adopted by Goutaki *et al* (40), studies with an unclear reporting of diagnostic methods or incomplete set of diagnostic tests performed were not excluded, as the focus of this review was to describe the different outcome measures in use for PCD and highlight new indicators. Details of diagnostic data used by each of the included studies can be found on Appendix C. Studies with overlapping patient population were not excluded if they reported on different clinical measures.

We excluded studies that were not original research, conference abstracts where the manuscript was published later or that did not contain sufficient information on outcome measures, and full texts that were irretrievable. Manuscripts that reported on ENT symptoms only were also excluded, as a separate review on ENT measures has been planned.

6.2.3 Manuscript selection

The search was performed on 10th October 2017. Two reviewers (BR, CJ) independently assessed titles and abstracts for eligibility. Full text was obtained for all studies deemed relevant by the reviewers or where there was uncertainty or disagreements on eligibility. Where disagreements remained after full text review, the manuscripts were discussed with a third person (JSL). One reviewer (BR) manually searched the reference lists of all eligible studies for additional manuscripts.

We extracted data using a standardised data extraction form in Excel, defined a priori. The form was piloted on five randomly selected manuscripts and then refined. Two reviewers (BR, FG) extracted data for half of the eligible studies each on the following: publication details (authors, title, year of publication, country, and journal), study characteristics (data collection period, study design, countries that contributed with data, inclusion criteria, clinic type, sample size, population characteristics, and diagnostic data), and outcome details (outcomes reported, definitions used, equipment used and measurement details, and correlation between different outcome measures).

An additional third reviewer (CJ) extracted data from eight manuscripts, an overlap of approximately 10% of the total manuscripts included in the study, to ensure consistency on data extraction by the two main reviewers (BR, FG). This approach was taken due to the high volume of studies retrieved and extensive information extracted from each, hindering the original plan to double extract all data. BR and FG discussed in detail the full text and the data extracted from all

selected studies to confirm there was agreement on eligibility and on the information recorded in the data extraction form.

6.2.4 Definition of outcome measures and classification into subgroups

We defined outcome measures as any clinical measure that was used a) to monitor patients over time or b) as a marker of disease severity. For example, presence of mucoid *Pseudomonas aeruginosa* is correlated to worse clinical outcome in CF (203) and therefore studies that described rates of PCD patients infected by the pathogen were included, with colonisation by *Pseudomonas* considered a marker of disease severity.

Clinical outcome measures were classified as (i) study outcome, if defined and used as actual study outcome; or as (ii) population descriptor. The latter was applied to clinical measures that have been previously associated with a worse outcome in PCD or CF but were used to characterise the study population as opposed to actual study outcomes (*e.g.* baseline characteristics of FEV_1 measures in the PCD study population) or those that could potentially be used in future studies (*e.g.* cough frequency).

6.2.5 Statistical analysis

Continuous variables were reported as median and interquartile range, after checking for Normality. Categorical variables were reported as proportions. Figures were plotted in R and Tableau 2019 v4.0.

Reporting of results adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Extension for Scoping Reviews (PRISMA-ScR) checklist and recommendations (160). The review protocol was initially registered in PROSPERO (registration number: CRD42018051111), the international prospective register of systematic reviews, but had to be withdrawn as the registry does not accept scoping reviews. The full protocol is available upon request.

Box 1. Key terms used in the search strategy in Embase, with results from each search term

1. Exp kartagener syndrome/ (1220)
2. Exp ciliary motility disorders/ (2072)
3. primary ciliary dyskinesia.ti,ab. (1385)
4. 1 OR 2 OR 3 (3056)

5. exp respiratory function test/ or exp lung function test/ (87859)
6. exp vital capacity/ (9493)
7. exp spirometry/ (35581)
8. exp airway resistance/ (12324)
9. exp blood gas analysis/ (21614)
10. exp bronchial provocation test/ (3264)
11. capnometry/ or exp lung function test/ or patient monitoring/ (164554)
12. exp lung compliance/ (10444)
13. exp lung volume measurements/ (99050)
14. exp plethysmography, whole body/ (3283)
15. exp pulmonary gas exchange/ (10698)
16. Bronchiectasis.ti,ab. (10739)
17. exp bronchiectasis/co, di, dm, ep, et, pc, su [Complication, Diagnosis, Disease Management, Epidemiology, Etiology, Prevention, Surgery] (5077)
18. 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 (284811)

19. Outcome parameter$.mp. or Treatment outcome/ [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading] (727565)
20. exp hospital admission$/ or patient readmission/ or hospitalization/ (443673)
21. Hospital$.mp. (2117005)
22. mortality/ (775426)
23. morbidity/ (365260)
24. life expectancy/ (43058)
25. (Day$ antibiotic$ or antibiotic$ course$).mp. (1335)
26. Need for surgery.mp. (4001)
27. Quality of life/ (362178)
28. Disease progression (224730)
29. 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 (3775122)

30. Pulmonary exacerbation$.mp OR Disease exacerbation/ (58109)
31. Respiratory rate.mp (15545)
32. Monitoring, physiologic/ (1307)
33. Cough/ OR cough frequency.mp (33801)
34. Respiratory sounds/ OR respiratory frequency.mp OR breathing frequency.mp (9362)
35. Rhinomanometry/ OR exp Otorhinolaryngologic Surgical Procedures/ OR exp Otorhinolaryngologic Diseases/ (456842)
36. Sputum/ OR sputum clearance.mp OR sputum colo?r.mp (25673)
37. 30 or 31 or 32 or 33 or 34 or 35 or 36 (582840)

38. Tomography, Emission-Computed/ or tomography.mp. (903523)
39. Magnetic Resonance Imaging/ or MRI.mp. (502846)
40. Radiography/ or Xray.mp. or Radiography.mp. (539099)
41. Diagnostic Techniques, ontological/ OR Hearing tests/ OR Audiometry/ (23042)
42. Exp Otitis Media/ (32598)
43. Body mass index/ (225715)
44. Symptom score.mp OR symptom scale.mp (16609)
45. Inflammation/ or Inflammatory markers.mp or biomarkers/ (617588)
46. 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 (2482985)

47. 18 or 29 or 37 or 46 (6281627)
48. 4 AND 47 (2236)
49. limit 48 to (human and yr="1996 -Current") (1513)
50. remove duplicates from 49 (1453)

6.3 Results

Two thousand five hundred and sixty-two abstracts were identified, of which 2134 were reviewed after exclusion of 428 duplicates. Ninety-four manuscripts were reviewed in full, of which 71 met the inclusion criteria and were therefore included (Figure 15).

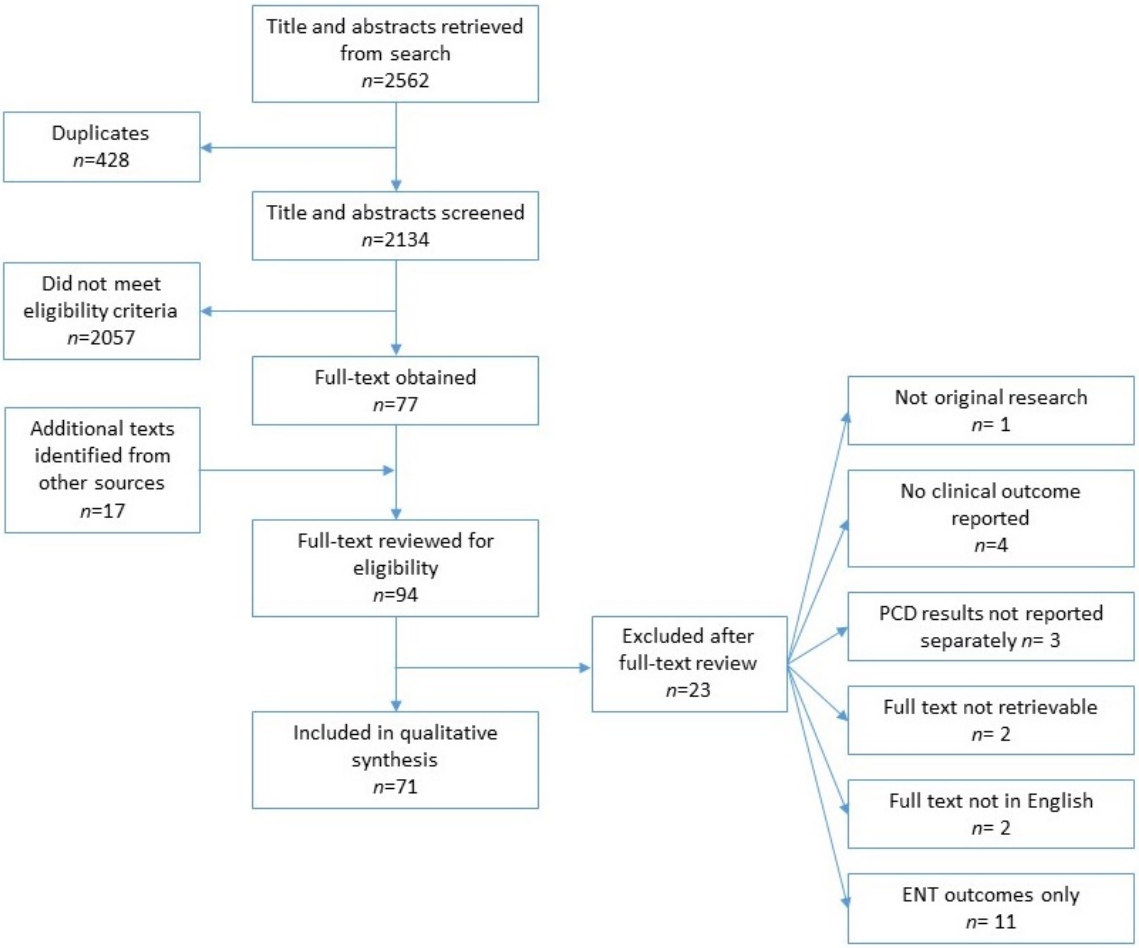


Figure 15. PRISMA flow diagram for the selection of studies reviewed and included in the systematic review.

6.3.1 Study characteristics

The manuscripts included data on 3373 patients with PCD, with a median of 29 PCD patients per manuscript (range 10 to 217). Studies contained data collected from 18 different countries, however only 10% presented data from more than one centre (Figure 16). Publications with a higher number of PCD patients were from multi-centre studies, with the largest study containing data from Israel, Belgium, Germany and Italy (204). There were two single-centred studies with over 150 PCD patients, both retrospective cohorts spanning a period of 20 years (84, 121). The number of larger multicentre studies has increased in recent years, highlighting the role of PCD networks such as BESTCilia, BEAT-PCD and the Genetic Disorders of Mucociliary Clearance Consortium in advancing collaborative research in the field (194, 205, 206).

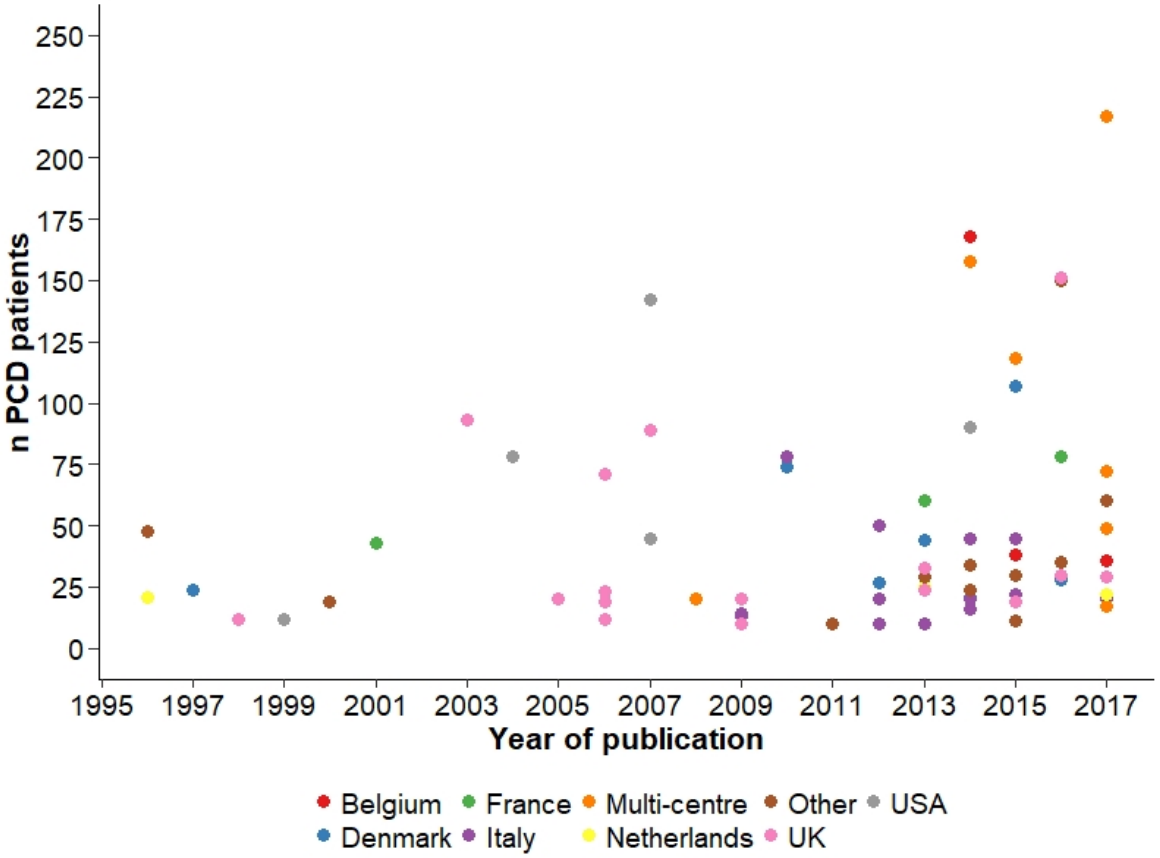


Figure 16. Number of PCD patients per study by year of publication, stratified by country (colours). Orange circles represent manuscripts that reported on data from more than one country.

6.3.2 Outcome measures reported

Sixty-four studies reported on a total of 22 study outcomes, while an additional seven only presented population descriptors (Figure 17). Forty-seven reported on both study outcomes and population descriptors and 17 reported exclusively on study outcomes (Appendix C).

Spirometry-derived parameters were the most frequently reported clinical outcome measures, followed by HRCT. Microbiology and anthropometric measures were more often reported as descriptors than as study outcomes (Figure 17). Fertility was reported only as population descriptor, with studies presenting data on sperm motility, subfertility, number of ectopic pregnancies and miscarriages, number of children born without the use of assisted reproductive technologies, births after intracytoplasmic sperm injection and births from in vitro fertilisation.

In the following sections we highlight the most frequently used clinical outcome measures, which included lung function measurements, imaging, microbiology, anthropometry, and HRQoL. Some

individual studies are highlighted throughout for illustrative purposes or because they represent significant developments in PCD research.

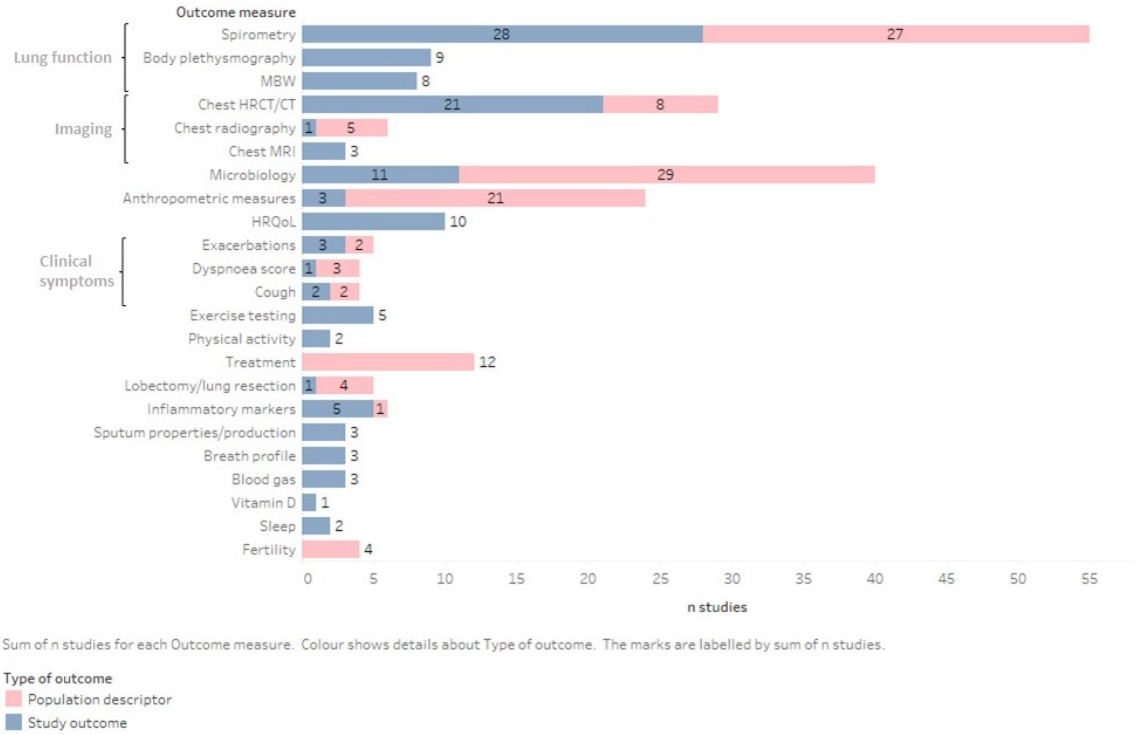


Figure 17. Number of studies that reported outcome measures in PCD as either study outcome or population descriptor. Studies often reported on more than one outcome measure and might therefore be featured multiple times.

6.3.3 Standardised definitions

Definitions of outcome measures varied considerably between studies. For example, definition of chronic colonisation differed in terms of requisite number of positive cultures (two, three or at least four) over various time periods (within six months, at least six months or within a year). Sampling frequency of sputum and other microbiological specimens also varied between studies. Some studies did not record whether patients had a pulmonary exacerbation at the time of sampling. These differences likely biased results, particularly when reporting prevalence of different respiratory pathogens (Figure 18).

Of ten studies reporting on microbiology as study outcomes, eight provided data on chronic colonisation by potentially pathogenic bacteria (105, 121, 130, 207-210). The terms chronic colonisation and chronic infection were sometimes used interchangeably, and the classification applied to define them varied. Two studies used the Leeds CF criteria (or a modified version) (130, 204), two the European consensus for antibiotic therapy against *Pseudomonas aeruginosa* in CF (208, 211) and one the Copenhagen criteria (209).

Details on definitions for all clinical outcome measures reported as study outcome or population descriptor are provided in Appendix C.

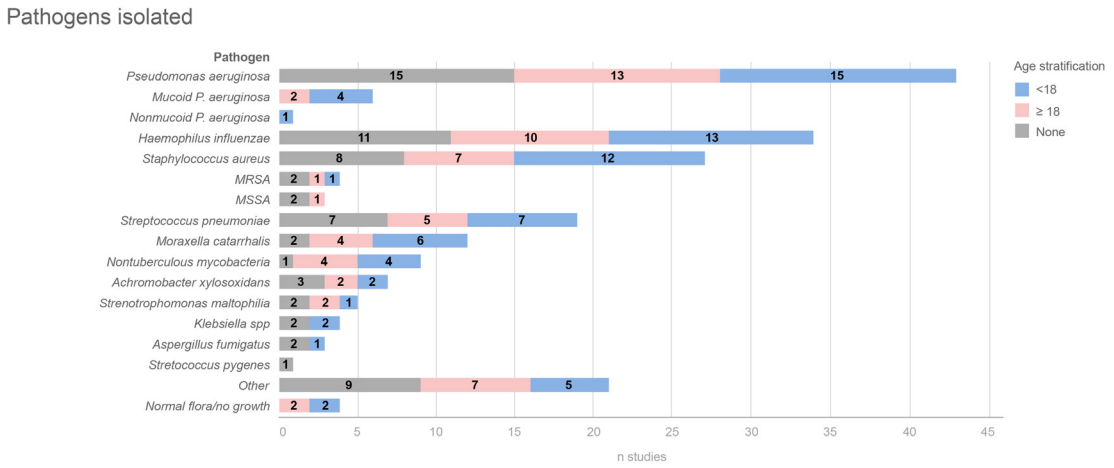


Figure 18. Number of studies that reported on each pathogen, stratified by age group (<18 and ≥18 years of age, or not differentiated).. The "Other" category includes less frequently reported pathogens such as *Burkholderia cepacian*, *Candida albicans*, *Serratia mercescens*.

6.3.4 Lung function

6.3.4.1 Spirometry

Of the 55 studies describing spirometry data, only 24 (43.6%) reported adherence to the ERS/ATS guidelines despite the existance of clear guidelines on standardisation of measurement and reporting of spirometric indexes (123). Eleven studies (20%) presented data on FEV_1 z-scores, eight as a study outcome. FEV_1 % predicted was more often reported by studies using the measurement as a descriptor (*n*=23 *vs* 21 as study outcomes). One study provided raw values of FEV_1 (212). Studies reporting on FEV_1 z-scores were published more recently, from 2012 onwards (Figure 19).

Two studies had a long follow-up period but were retrospective. One had a median follow-up of 9.5 years (range 1.5 to 30.2 years) and reported a high degree of heterogeneity in the course of lung function after diagnosis. FEV_1 and FVC were not correlated with age at diagnosis or baseline lung function (120). The other study followed adult patients for a median of 7 years (range 1 to 34 years) and found that FEV_1 declined 0.49% predicted annually. This decline was steeper in patients with MTD defects, compared to those with ODA or combined ODA and IDA defects by TEM, and compared to those with normal or inconclusive findings on TEM (121).

Four studies compared FEV_1 before and after the use of bronchodilators. Studies used inhaled salbutamol (213, 214) or albuterol (215) to assess reversibility and one study performed methacholine challenge before and after the use of salbutamol compared to placebo (108) (see 6.3.9). The studies found no evidence that the routine use of bronchodilators was beneficial in PCD patients.

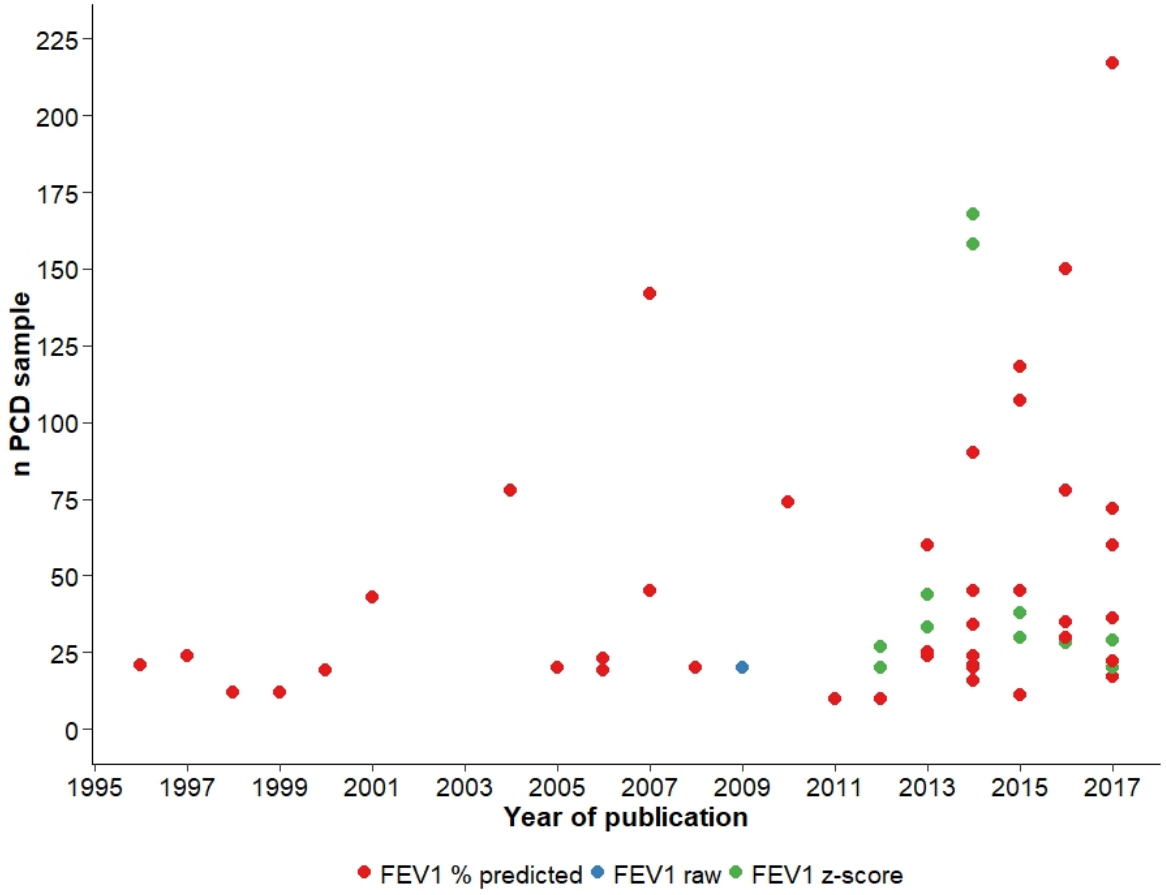


Figure 19. Studies that presented FEV_1 as study outcome or population descriptor . (*n*=55). Each circle represents one study, coloured by the spirometric index reported.

6.3.4.2 Body plethysmography

Nine studies (12.7% of the studies included in this review) reported on body plethysmography parameters as study outcome (Appendix C). Lung residual volume % predicted was reported by all studies, total lung capacity % predicted in four studies, and the remaining 19 parameters were scarcely reported. Three of these studies additionally measured FVC and FEV_1 using plethysmography devices, which also have the option for spirometry.

6.3.4.3 Multiple breath washouts

Eight studies (11.3%) reported on parameters derived from MBW test as study outcomes (Appendix C). LCI was the most frequently reported parameter. Seven studies (87.5%) presented

standardised indexes (z-scores) for LCI. Four (50%) presented values for S_{cond} and S_{acin} z-scores, which represent ventilation inhomogeneity in small *conducting* and *acinar* airways respectively. However, studies used different inert tracer gases and equipment; half of them used 0.2% sulfur hexafluoride (SF_6) and the other four used nitrogen (N_2), limiting comparisons of results between studies.

6.3.5 Imaging

Thirty-two manuscripts (45.1%) reported on imaging outcomes, with 21 presenting them as study outcome and 11 as population descriptor (Appendix C). Of the latter, four studies (36.4%) had spirometry measures as study outcome and only provided information on presence or absence of bronchiectasis, diagnosed through chest HRCT, CT, or radiography. The remaining studies did not report on specific outcomes, with data only on descriptors including spirometry, microbiology, growth, and fertility.

Four studies reported on both chest radiographs and HRCT, and a further two studies on chest HRCT and magnetic resonance imaging (MRI).

6.3.5.1 Chest Radiography

Bronchiectasis, seen on chest radiography, was used in one study as study outcome and in five as population descriptor (Appendix C). There are no PCD-specific radiography scoring systems, so studies used different scales to report findings. For example, Jain *et al* (216) used a modified version of the Chrispin-Norman score, which was originally developed for CF (217). The authors reported hyperinflation, bronchial wall thickening, atelectasis and consolidation as the most common features seen in PCD patients. Kennedy *et al* (119) developed a study-specific score for bronchiectasis severity, which was higher in patients that were older, had compromised lung function and had cultured mucoid *Pseudomonas aeruginosa*.

6.3.5.2 Chest HRCT

Chest HRCT and/or CT was used as study outcome in 21 studies (29.6%), with an additional eight studies reporting it as population descriptor (Figure 17 and Appendix C). Unsurprisingly, all 20 studies that presented information on their scoring system reported on bronchiectasis. Airway wall thickening and mucous plugging were the second most common features described, in 16 studies each.

Studies adopted modified versions of different scoring systems as there are no PCD-specific ones available. Seven studies used modifications of the Brody score (218), four used the Bhalla score (or

modified versions) (219), another four applied the Helbich score (or modified versions) (220) and a further five used other systems. Of the latter, three studies used a study-specific score, one by combining the Brody and Bhalla scores (221). Table 23 provides detailed information on each of the CF-derived scales.

The use of different measurement scales resulted in inconsistent reporting of sub-scores (Figure 20). For example, studies measured extent of bronchiectasis in various ways: a) number of bronchopulmonary segments affected, b) percentages of each lobe involved, c) scores from 0 to 3, d) percentages of central lung and peripheral lung involvement, or e) size of largest and average bronchopulmonary segment involved. Additionally, not all studies using the Brody score reported on the same sub-score components, likely due to study-specific modifications (Figure 20). Mucus plugging sub-scores were derived from the size of the plug (*i.e* small, large), location of the plug (*i.e* largest airways, small airways, peripheral lung, and central lung for Brody score; or number of segments for Helbich score) or using a mucus classification score (Bhalla score). In another study that used a modified combination of the Brody and Bhalla scores, the partition of lungs into different segments followed a regional approach as opposed to the commonly used pulmonary segmentation approach to expedite the time needed for scoring each CT scan in routine clinical practice (221).

Two studies comparing CT scores in PCD and CF patients found no significant difference in the global Brody score (105, 111). However, Maglione *et al* (111) reported a significantly higher sub-score for severity of collapse or consolidation in PCD compared to CF, while Cohen-Cymberknoh *et al* (105) found that the lower lobes were more frequently affected in PCD compared to the typical upper lobes compromise seen in CF (28, 105, 119, 121).

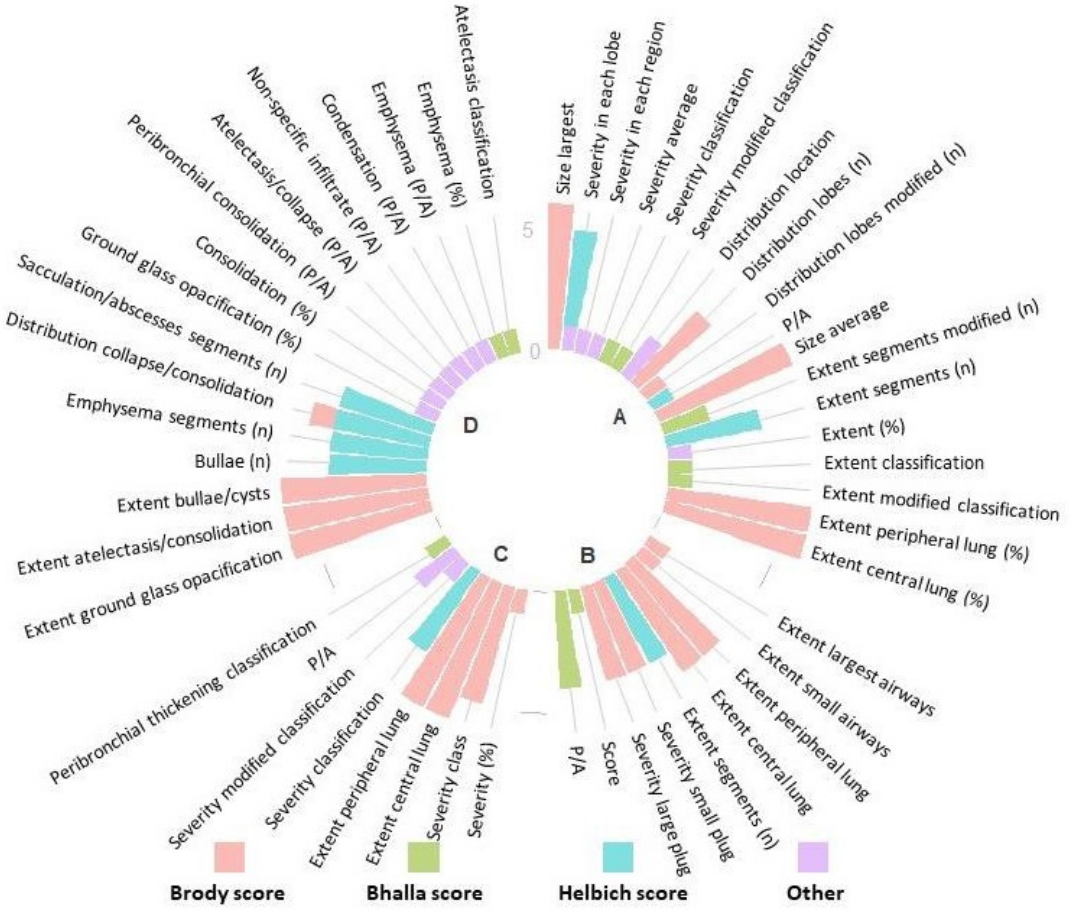


Figure 20. High-resolution computed tomography (HRCT) and CT outcome measures from 20 studies that reported on HRCT/CT scans as study outcome.. Sub-scores are displayed as A bronchiectasis, B mucus plugging, C airway wall thickening, and D parenchyma.

6.3.5.3 Chest MRI

Only three studies (4.2%) reported on chest MRI as study outcome (Appendix C). All studies that presented chest MRI data applied a modified Helbich scoring system. When examining sub-scores, the only significant difference found was a higher sub-score for severity of collapse and consolidation. The study compared 20 PCD patients to 20 patients with mild CF, with no difference seen between total MRI scores and other sub-scores (208).

6.3.6 Microbiology

Forty-four studies (62%) reported on microbiology, 11 as study outcomes and 29 as population descriptors (Appendix C). The most commonly described respiratory pathogens were *Haemophilus influenzae* and *Pseudomonas aeruginosa*, followed by *Staphylococcus aureus* (Figure 18) (28, 105, 121, 126, 130, 204, 208, 211). Some studies distinguished between mucoid and non-

mucoid strains of *P. aeruginosa* while others simply reported on *Pseudomonas* infection. Similarly, *S. aureus* subtypes were inconsistently stratified across studies, with some reporting methicillin sensitive (MSSA) and resistant (MRSA) strains separately.

Not all studies stratified pathogen prevalence by age group (Figure 18). Since the prevalence of bacterial species is known to change with age (130, 180, 222), future studies should present a breakdown of pathogens by age group.

Interpretation of pathogen prevalence was limited by the lack of a universal panel that could be applied consistently across different centres. Rogers *et al* (223) highlighted that some of the dominant genera of bacteria found in the sputum of PCD patients were from those unlikely to be detected without specific growth conditions being present. Variations in the frequency of specimen collection and type of specimen (*e.g.* sputum, bronchoalveolar lavage) can also affect pathogen prevalence.

6.3.7 Other outcome measures

6.3.7.1 Anthropometry

Only three studies (4.2%) reported anthropometric measures as study outcomes, of which two used z-scores and one reported on body mass index (BMI) percentiles (supplementary table 2). A further 21 studies (29.6%) described height, weight and/or BMI as population descriptor, highlighting the availability of these data.

6.3.7.2 Health-related quality of life scores

Ten studies (14.1%) reported on HRQoL as study outcomes, all of which were published before QoL-PCD was validated (198, 224, 225) and therefore did not use the disease-specific instrument. The SGRQ was the most frequently used tool (*n*=8), followed by the 36-item short form survey (SF-36) in six studies (Appendix C).

6.3.7.3 Pulmonary exacerbations

Five studies (7.0%) reported on pulmonary exacerbations, three as study outcomes. However, definitions of pulmonary exacerbations varied between studies. The only RCT in PCD to use it as a primary outcome defined an exacerbation as respiratory symptoms that led to initiation of systematic antibiotic treatment irrespective of culture results, or a decline of at least 10% in FEV_1 % predicted compared to baseline at screening and randomisation (106) (See 6.3.9). Joensen *et al* (209) applied a definition developed for CF studies (226) (Appendix C).

Ratjen *et al* (126) studied a subset of patients that experienced an episode of exacerbation, defined as an increase in lower airway symptoms treated with oral antibiotics, and Sunther *et al* (125) only included patients with pulmonary exacerbation, defined as change in respiratory status for which intravenous antibiotics were needed.

6.3.8 Comparison between outcome measures

Most studies comparing two or more outcome measures used spirometry as the reference to which other outcomes were compared (Figure 21). Eight studies describing imaging modalities reported on agreements or correlations with other outcome measures (111, 117-119, 214, 227-229). The most common comparison was between spirometry derived FEV_1 and HRCT, with studies presenting contradictory findings. Two studies (119, 228) found an agreement between the two outcomes, despite one using a modified Bhalla system and the other a study-specific scoring system. The other four studies (111, 214, 227, 230) reported no association.

Studies comparing HRCT to indices derived from body plethysmography, chest MRI and microbiology found significant correlations. However, these were generally limited to specific sub-scores (*e.g.* bronchiectasis for body plethysmography, and collapse and consolidation for MRI) as opposed to the global score.

FEV_1 was compared to MBW-derived LCI in five studies, also with contradictory results. Two studies reported no association (112, 230), while the other three (117, 118, 231) found correlations between some parameters. Boon *et al* (117) found a significant negative correlation between LCI and FEV_1, FEV_1/FVC ratio and FEF_{25-75} z-scores. They also reported that LCI z-scores were concordant with total CFCT scores, a variant of the Brody score, in 83% of the patients, while Irving *et al* (118) found no correlation. Green *et al* (231) found no correlation between LCI and FEV_1 z-scores in PCD patients but reported significant correlation between $LCI_{2.5\%}$ and FEV_1/FVC ratio and FEF_{25-75} z-scores; the latter was also reported by Irving *et al* (118). Differences in tracer gas, equipment used for MBW measurements, inclusion criteria, age groups, and small samples sizes could explain inconsistent findings.

LCI might be more sensitive to detect early or mild disease as five patients (15.2%) had abnormal LCI but normal FEV_1 z-scores (118). LCI was also shown to be more sensitive than FEV_1 to detect lung structure abnormalities (117).

Other associations between outcome measures are shown in Figure 21.

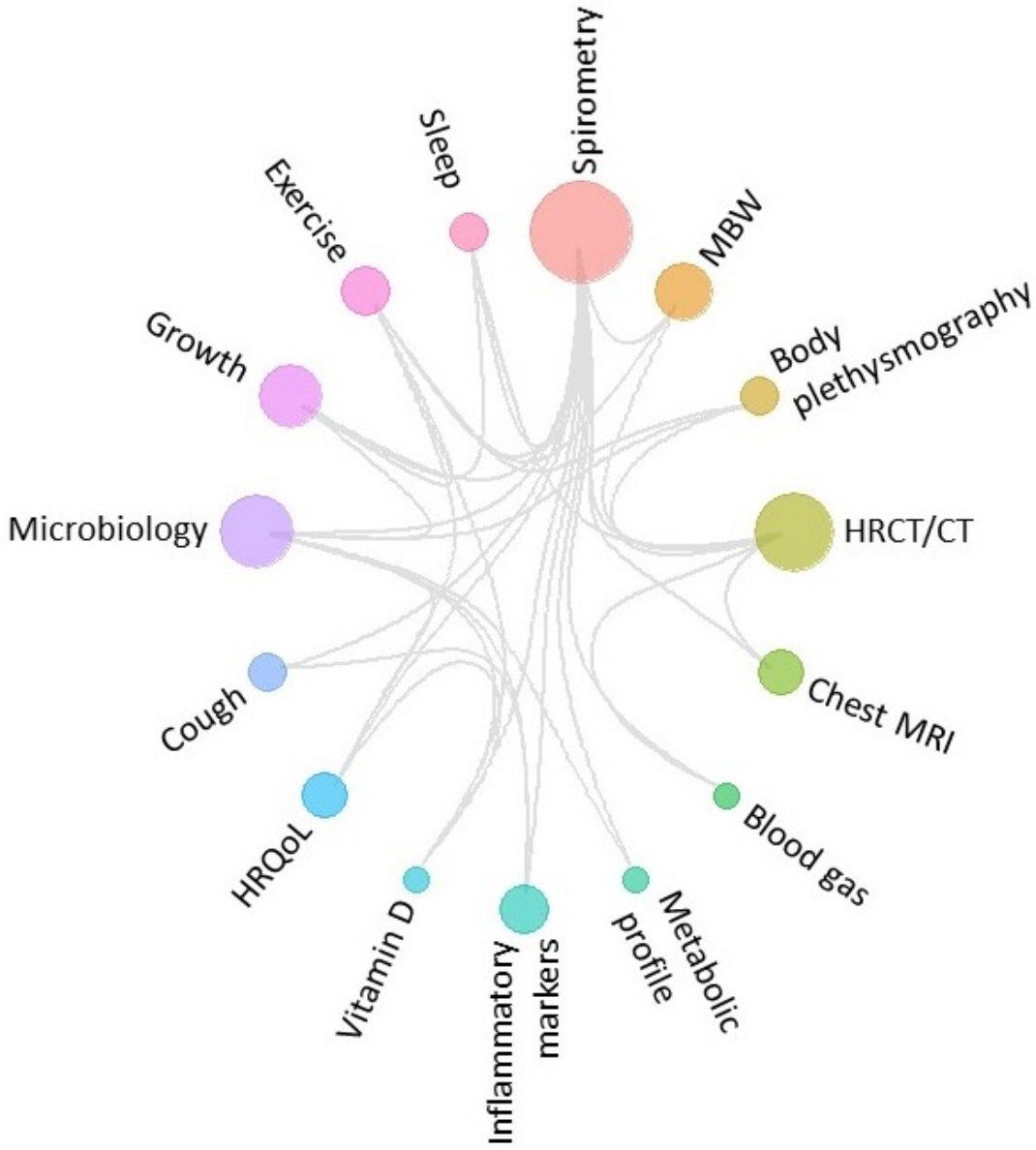


Figure 21. Correlations or associations between outcomes measures.. Connections between circles depicts the correlated outcomes, with the size of each circle representing the number of studies that reported correlations or associations of that particular outcome.

6.3.9 Randomised controlled trials

Only four of the included studies were RCTs, all adopting a crossover design (Table 10). The most recently published study was a 28-week double-blind crossover RCT with a wash-out period of four weeks. The authors investigated the effect of hypertonic saline on HRQoL in PCD adults, measured by the SGRQ and QOL-B, compared to isotonic saline (106). They observed a significant

improvement on the Health Perception scale of QOL-B but not on the total score or sub-scores of the SGRQ, or any of the secondary outcomes.

Gokdemir *et al* (232) assessed spirometry measurements (FEV_1, FVC, PEF and FEF_{25-75} % predicted) in PCD children using two different airway clearance methods. Half performed conventional pulmonary rehabilitation for five days in hospital followed by a two-day wash-out period and then high frequency chest wall oscillation for another five days at home. The authors reported no significant difference in lung function parameters between the two groups; however, techniques differed between the settings. Another crossover RCT found no significant difference in FEV_1 % predicted and in bronchial hyperresponsiveness after the use of salbutamol or placebo in PCD children at both three and six weeks compared to pre-treatment measurements (108).

Noone *et al* (233) found significant increase in mean whole-lung clearance rates of a radionucleotide marker after inhalation of uridine-5'-triphosphate compared to placebo during a series of controlled coughs to induce mucociliary clearance in PCD adolescents and adults. This difference remained significant after 60 and 120 minutes from inhalation; however, particle retention at 24 hours was the same in both groups.

Table 10. Summary of study characteristics of crossover randomised controlled trials included in this scoping review.

Authors (year of publication)	***n* PCD patients**	**Intervention**	**Reference group**	**Limitations**
Paff *et al* (2017) (106)	22	Hypertonic saline	Isotonic saline	Small sample size; Non-disease-specific outcomes; Isotonic saline might have beneficial effect.
Gokdemir *et al* (2014) (232)	24	High frequency chest wall oscillation	Conventional pulmonary rehabilitation	Small sample size; Short follow-up and wash-out periods; No *a priori* definition of clinically significant effect.

Authors (year of publication)	***n* PCD patients**	**Intervention**	**Reference group**	**Limitations**
Koh *et al* (1999) (108)	19	Salbutamol	Placebo	Small sample size; Over 80% had bronchiectasis (disease severity); Unclear if all had PCD (only 42% had 'hallmark' TEM); Lack of definition for clinical stability.
Noone *et al* (1999) (233)	12	Aerolised uridine-5'-triphosphate	Placebo (0.12% saline)	Small sample size; All had bronchiectasis (disease severity); Unclear clinical significance as differences were only temporary.

6.4 Discussion

This scoping review identified 22 different clinical outcome measures that have been used in PCD research. We found a high degree of heterogeneity in the definitions of outcome measures. Comparison of results from studies reporting on chronic colonisation by respiratory pathogens and pulmonary exacerbations were limited by the lack of a standardised definitions. Spirometry and chest HRCT were the clinical outcome measures most frequently reported as study outcome.

Spirometry is widely available, easy to perform and does not require expensive equipment (123, 191, 234); however, researchers have questioned its appropriateness as a surrogate measure to monitor disease progression in PCD (111, 125, 235). Flow and volume measured by spirometry are affected by resistance of the larger airways and therefore spirometry measurements are more sensitive to detect proximal airway disease. A meta-analysis found that mean FEV_1 ranged from 51% to 96% predicted, with high heterogeneity between studies that could not be explained by

age or other explanatory factors (see 2.5.1) (113). Studies that did not report on the reference values used or did not provide information on quality control had lower mean FEV_1, underscoring the importance of standardising the measurement and reporting of spirometry-derived parameters. Clinical status at the time of measurement was rarely reported and therefore could not be included in the meta-regression. A recently published (after our search) large multicentre retrospective cohort (n=991) reported consistently low FEV_1 z-scores in patients with PCD compared to reference data, similar to the low levels seen in CF patients (73). More recently, Rubbo *et al* (6) showed that levels can be significantly lower in children with PCD compared to CF.

To our knowledge, no study has investigated the timing of physiotherapy in relation to spirometry, which is a significant limitation as airway clearance techniques can anecdotally improve spirometric indices by removing excess mucus from the respiratory tract. An ongoing multicentre prospective cohort is investigating variability of lung function in stable PCD patients, adjusting for factors such as timing of inhaled medication and respiratory physiotherapy (205, 206) (https://clinicaltrials.gov/ct2/show/NCT03704896).

MBW is thought to be more sensitive for monitoring early lung involvement in PCD as it measures peripheral airway involvement, with a recent study showing significant increase in LCI over the period of one year in 42 patients with PCD (236). However, a limitation of LCI is the long washout time and test-duration, which is particularly problematic for patients with compromised lung capacity and young children. Studies looking at shorter washout periods have shown promising results, with $LCI_{5\%}$ providing a good alternative to the more conventional $LCI_{2.5\%}$ (231, 237, 238).

Chest HRCT has been proposed as a surrogate outcome measure in the assessment of lung disease. However, there are no validated scoring systems for PCD, and radiation-low modalities are needed. All studies included in this review used CF-derived scoring systems (218-220), despite significant pathophysiological differences between the two conditions (104, 239).

Location, distribution and frequency of features seen in HRCT scans of patients with PCD differ from those with CF (239). The weights applied to each feature might not be suitable for PCD as CF-derived scoring systems do not reflect the range and severity of structural changes in PCD. Studies found that extensive tree-in-bud pattern of mucus plugging, bronchoceles or nodules, thickening of interlobar and interlobular septal, and atelectasis mostly seen as collapse of whole lobes were frequently described in PCD but uncommon in CF (239-241). Reporting only the overall CT scores might be misleading, particularly when using a non-disease-specific score. These findings underscore the need for disease-specific CT scoring systems (239). Recently, Hoang-Thi *et al* (242) developed an automated CT scoring for adults with PCD, which had moderate to good correlation with FEV_1 and FVC.

MRI scans of the chest have been considered of limited value due to intrinsic characteristics of the pulmonary tissue, and the presence of physiological motion resulting in poorer resolution and motion artefacts. Recent research has focused on improving techniques in order to obtain better quality images (243, 244). If this is achieved, MRI could be a good radiation-free alternative for longitudinal monitoring of structural and functional changes in PCD.

Lack of agreement between spirometry, chest HRCT, MBW and chest MRI parameters reported by some studies might reflect variations on measurement and reporting of outcomes. Discrepancies could be explained by different scoring systems for HRCT, inability of some of the outcome measures to accurately monitor lung disease progression in PCD, or true variability between populations (*e.g.* underlying genetics, differences in disease severity or treatment). Interpretation of findings was limited by the retrospective nature of most studies. In some cases there was a significant time lag between the measurements performed using the compared methods (117), or tests were applied to different sub-populations (*e.g.* HRCT scans conducted only in the older population with more severe lung disease (118) or conducted at different timepoints of clinical stability (111)). Contradictory results could also be attributed to variations in study design or to underpowered studies leading to insignificant results.

A recently published (235) study comparing spirometry, MBW and structural and functional chest MRI found that over half of the patients with abnormal LCI values and MRI scores had normal FEV_1 z-scores. LCI was not able to distinguish between reversible and irreversible lung damage, despite being more sensitive than spirometry to detect changes. Combining different modalities (*e.g.* MRI and MBW) is therefore necessary to accurately capture changes in the lungs of PCD patients.

6.4.1 Limitations

Our review was limited by the quality of the information provided in the studies. Small sample sizes were a common limitation in most studies, highlighting the importance of national and international disease registries, large collaborative multicentred studies and standardised definitions that enable pooling of data (55, 94, 95, 245). Few studies included sample size calculations, hampering the interpretation of statistically insignificant results due to underpowered samples.

The search was not updated after October 2017 as it would have been an enormous undertaking and we had limited time and resources. However, we have contrasted and compared our findings with some of the most relevant recently published studies in the discussion. As our aim was to identify the evidence available and describe definitions for clinical outcome measures used in PCD

research, we opted to conduct a scoping review and therefore we did not critically appraise the studies included in this review. We did not perform quantitative analysis as studies were heterogeneous, impeding a formal meta-analysis to be carried out. In fact, one of the aims of this review was to highlight this heterogeneity.

RCTs and prospective cohort studies with long follow-up periods are uncommon in rare diseases due to the small sample sizes available, high costs and limited commercial interest from pharmaceutical industries (5, 91, 246) (see 2.1.1). As a result, the majority of PCD studies are cross-sectional, case-controls or small cohort studies with limited follow-up. Interventional studies require close international collaborations and data sharing, and the success of these will depend on the selection of appropriate outcome measures.

6.4.2 Recommendations

We advocate that outcome measures for use in future prospective trials must fulfil the following criteria: a) be measured across different studies in a standardised manner, using the same definitions; b) be consistently reported by a sufficient number of longitudinal studies; c) use currently recommended standardised measures (*e.g.* z-scores based on GLI recommendations) and d) be embedded within the current knowledge of PCD pathophysiology and natural history. This will require a consensus statement.

Most of the currently used endpoints and outcome measures have been extrapolated from CF literature. Even though this is an adequate starting point, disease-specific and validated outcome measures are needed. This scoping review provides the basis for the development of an expert-led consensus statement, which should incorporate the views of all those involved in service delivery and those affected by the disease, such as clinicians, healthcare professionals and patients and family members of patients with PCD.

Our review found that spirometry and chest HRCT were the most frequently used outcome measures to monitor PCD patients. Differences in how these outcomes were measures and reported impeded us from formally assessing them in a meta-analysis; however, there is sufficient evidence to recommend the continued use of these in PCD research, particularly if disease-specific scoring systems were to be developed and validated for use in PCD. Additionally, the existing disease-specific outcome measures, QOL-PCD and the expert-led pulmonary exacerbation consensus, should be prioritised as main or secondary outcome measures in future prospective studies (Table 11).

Spirometry was the most frequently outcome used for disease monitoring but there were major problems with standardisation on the measurement and reporting of FEV_1. Powered well-designed studies are needed to investigate the suitability of spirometry-derived parameters as accurate and sensitive surrogate markers in PCD.

HRCT might be a good candidate for longitudinal follow-up of lung disease progression in PCD, particularly modalities using low radiation (247, 248). However, a disease-specific scoring system must be developed and validated for use in PCD. Agreement between HRCT and other outcomes were limited to sub-scores as opposed to global score, emphasising the need for PCD-specific scores that consider the distribution, frequency and patterns of lung compromise in this population, and that can be easily applied by clinicians without being unnecessarily time-consuming.

The Food and Drug Administration (FDA) and the European Medicines Agency (EMA) encourage the use of patient-centred outcome measures in RCTs. A systematic review on the patient's experience of PCD reported worsening of respiratory symptoms with age, which was also associated with decline in the physical and mental domains (199). QOL-PCD is the only validated disease- and age-specific cross-cultural outcome measure in PCD (198, 224, 225, 249). Importantly, this HRQoL instrument has been translated into several languages, providing standardised patient-reported outcome measures for use in various countries (250-253). QOL-PCD correlated well with the Sino-Nasal Outcome Test (SNOT-20) for upper airway symptoms, SGRQ-C for lower airways symptoms and SF-36 for physical functioning, role functioning and mental health (225).

The recently developed pulmonary exacerbation consensus established well-defined criteria to define respiratory exacerbations in PCD patients (see 2.5.1). The importance of disease-specific definitions was demonstrated in our review by the fact that all three studies reporting on lower airways exacerbations as study outcome used different definitions (15, 125, 126). There is now a need to validate the expert consensus and develop a separate definition for upper airway exacerbations.

Other clinical outcome measures highlighted in this scoping review (e.g. MBW, chest MRI, growth, microbiology) should be considered, depending on study design, data availability, and outcomes of interest. Equally, new clinical outcome measures might emerge and will need to be assessed. These will need to be revisited at a later date, when further evidence on their validity, reliability and sensitivity become available.

A working group composed of physicians from various specialties (*e.g.* paediatric pulmonologists, adult pulmonologists, cardiologists, ENT surgeons, neonatologists), healthcare professionals (*e.g.* nurses, physiotherapists and other allied healthcare professionals), and patients and their family members should be established. Additionally, our findings and recommendations can be applied to other diseases where there is a lack of a core set of outcome measures or of clear definitions for outcomes. This is particularly pertinent for rare diseases in general, as the number of patients available for prospective cohorts and RCTs is limited.

Table 11. Summary of recommendations on clinical outcome measures for use in PCD research.

Clinical outcome measure	**Strengths**	**Limitations**	**Future directions**
Spirometry	Routinely measured; Reported in 77.5% studies included	Unknown accuracy and sensitivity as surrogate marker for lung disease; Unstandardised reporting of indices	Investigate appropriateness of spirometry to monitor disease progression; Standardised reporting of spirometric indices (i.e. use of z-scores); Perform quality control
High-resolution computed tomography	Can assess structural lung damage; Reported in 36.6% studies included	Use of CF scoring systems; Frequent doses of radiation if used routinely for disease monitoring	Develop and validate PCD-specific scoring system
Health-related quality of life	Patient-centred outcome measure; QOL-PCD was developed and validated for use in PCD	Lack of a minimal clinically relevant difference	Adopt QOL-PCD as outcome measure in prospective longitudinal studies; Use translated and cultural validated versions, where available

Clinical outcome measure	**Strengths**	**Limitations**	**Future directions**
			Calculate minimal clinically relevant difference
Pulmonary exacerbations	Developed for use in PCD	Has not been validated	Validate definition; Use in future prospective studies

6.5 Conclusions

This scoping review highlights the variety of outcomes and definitions used in PCD research. It also underscores significant differences in measurement and reporting of outcomes. The absence of a core set of outcomes that are consistently reported across studies and the lack of standardised definitions hampers comparisons and needs to be urgently addressed through the development of consensus statements.

Validated disease-specific clinical outcome measures are needed to monitor disease progression in PCD in future prospective cohort studies and clinical trials. Until these are available, new studies should use outcomes that have been previously used in PCD research, such as those highlighted and recommended by this scoping review, to build up the body of evidence needed to meaningfully compare these, as well as test new outcomes that can be useful for regular monitoring.

Chapter 7 Comparison of service delivery models for PCD

7.1 Introduction

Health service delivery and organisational research investigates health services or systems as opposed to the state of health of individuals or populations, which is the focus of other disciplines such as epidemiology. Contextual evidence and underlying processes are paramount to understand how outcome measures can result in patient benefit. Context can explain variations between countries and settings, such as the provision of patient access to effective clinical care, frequency of monitoring, and treatment options. For example, CF studies have found that the development and implementation of specialised care facilities led to improvements in clinical care (254, 255).

Context can allow for a deeper understanding of research findings and their generalisability, which in turn informs the appropriateness of applying these to different healthcare systems. Differences detected between countries may reflect variations in the organisation of healthcare systems or demographics factors. In Switzerland, for example, insurance cover for PCD patients falls under birth defects and therefore diagnosis must be based on biopsied samples rather than on clinical symptoms or HSVA, reinforcing the concept that the choice of diagnostic methods can reflect underlying funding structures (4).

7.1.1 Literature review

Kuehni *et al* (4) and Strippoli *et al* (256) provided the most comprehensive studies of disparities in diagnostic and management services for PCD across Europe. The two manuscripts were derived from a single international survey conducted between January 2007 and January 2009, preceding all consensus and evidence-based guidelines. Using a two-stage design, the authors identified national representatives for each European country using the ERS membership roster. The representatives were responsible for reporting the number of tertiary centres in the country and for distributing the survey to every institution likely to be managing PCD patients, including tertiary, secondary (regional) and primary (defined as paediatric centres or small hospitals) care centres. As discussed in 2.3, the survey showed that the European region (adapted from the 2005 United Nations definition (257)) and the number of patients treated by each centre were important determinants of both age at diagnosis and use of recommended diagnostic methods (4,

256). Age at diagnosis was significantly lower in countries with higher levels of national government health expenditure (4). Higher health expenditure was also correlated to higher likelihood of using evidence-based diagnostic tests such as nNO, ciliary function, TEM, cell culture and genetic testing (256).

Only three countries (Cyprus, Denmark and Hungary) provided centralised PCD care at the time. The remaining had a decentralised system, with a mean of only four patients per centre. Management strategies were found to be highly heterogeneous and did not follow international recommendations. Northern and southern European countries used inhaled steroids routinely, despite lack of evidence that these are effective in PCD patients. Equally, dornase alfa was more commonly prescribed in larger centres and in the British Isles. On the other hand, airway clearance was prescribed more frequently in centres that managed a higher number of PCD patients and less frequently in Eastern Europe. Antibiotics were more likely to be used to treat exacerbations in Eastern Europe, while routine use of prophylactic antibiotics was more common in the UK and Ireland (256).

However, the survey had significant limitations. Only centres that cared for children were approached, therefore excluding the PCD adult population. Additionally, differences in healthcare structure meant that countries with centralised care only provided data from tertiary centres, which could have biased results. This is particularly relevant as the definition of tertiary care varies across Europe, limiting the generalisability of the study's findings. Data were derived from questionnaires sent to clinicians and therefore might not represent true clinical practice.

A recent meta-analysis found high heterogeneity in studies reporting on numbers of PCD patients from consecutive referrals to diagnostic services, which was attributed to different combinations of diagnostic tests to reach a diagnosis and variations in referral patterns between studies (258). The authors highlighted that referrals were still influenced by clinicians' awareness of PCD and personal experience despite the existence of guidelines on who should be referred for testing (25, 58).

7.1.2 Gap in the literature

Anecdotal evidence from conversations with members of other groups involved in research and clinical care in PCD suggest that significant differences between and within countries remain. However, these discrepancies need to be formally revisited, as data from the Europe-wide survey were collected more than ten years ago and a lot has changed in PCD care in recent years (25, 65, 103), particularly with the advent of wide-range PCD networks such as BEAT-PCD, fomenting research and knowledge exchange within the PCD community. Equally, the publication of the

evidence-based ERS PCD diagnostic guidelines in 2017 and attempts to develop standardised reporting of diagnostic tests have provided the basis for continued advancements in PCD diagnostics internationally (25, 67, 72, 85).

The English PCD national service is the most described PCD service delivery model in the literature (see 3.2). Little has been published on other service delivery models; mainly brief descriptions of the service in the methodology section of basic science or clinical research studies (120, 259).

7.1.3 Study aims

To describe the current state of delivery of care for PCD in different countries and to underscore changes in service delivery by comparing our findings with results from a survey conducted a decade ago (4).

7.2 Methodology

7.2.1 Mixed methods and theoretical approach

Quantitative research computes inputs and outputs without considering the underlying processes or the social context within which the research is conducted, and change occurs. In qualitative research, context is not considered a confounding factor that needs to be adjusted for, but rather an integral part of the shape, form and texture of a service delivery model, allowing for richer understanding of the structure, processes and outcomes, and the generalisability of findings to other models, in similar or differing contexts.

Paradigm defines the researcher's 'world view' or the assumptions adopted by the researcher when reviewing and analysing qualitative data. Positivism dictates that there are 'truthful facts' and it is the researcher's duty to uncover these. In contrast, the constructivist (or interpretivist) approach states that individuals have their own construction of 'social reality' and therefore there is not one truth but a truth to each different point of view. Our study design is better suited for this approach. One's reality is influenced by their point of view and experiences, which are imbedded in the participants' context and social web of interpretation and re-interpretation of experiences and opinions. Context, in this case, comprises of 'external factors' such as their country's healthcare and funding structures, and 'internal factors' determined by their own cultural background and experiences.

Equally, it is important to acknowledge that the researcher's point of view is also affected by their background and social-cultural experiences. The researcher examines the 'studied world' under

their adopted paradigm and theoretical perspective, which might differ significantly from observations of the same topic by other researchers with different backgrounds.

The broad definition of thematic analysis is "a method for identifying, analysing, and reporting patterns (themes) within data" (260). Thematic analysis is a process for qualitative analyses rather than a specific method. It is, therefore, not bounded by a particular theoretical or epistemological position and therefore can be used across a range of approaches (261). It offers a more flexible form of data analysis compared to classical methods such as grounded theory. However, qualitative analysis is an active process; the researcher identifies, selects, analyses and reports on themes and patterns. This underlines the need for researchers to acknowledge theoretical positions and values before making conclusions. While recognising the advantages of the flexibility provided by thematic analysis, Braun and Clarke developed a structured six-phase guide to applying thematic analysis (260). Despite losing some of its depth and complexity, this approach enables a rich overall description of the dataset, which was particularly useful in this project as it pertains to a greatly under-researched area.

7.2.2 Choice of data collection method

The selection of the appropriate method to collect qualitative data depends on the research topic and the question being addressed. Methods such as ethnography, where naturally occurring events are observed, are ideal to understand processes and structures; however, these are often time consuming, require extensive resources and carry ethical challenges. The implications in collecting data through ethnographic methods for this project would include travelling to several countries, recording days' worth of clinical activity, and obtaining ethical approval for observation of patient clinics from countries with varying degrees of patient and data confidentiality requirements.

Interviews are particularly useful to assess interviewees' attitude, values, interpretation of events, understandings, experiences, and activities, all of which are not easily obtained through other qualitative techniques (*e.g.* passive observation), or a more traditional quantitative approach (*e.g.* cross-sectional survey). Importantly, in order to obtain rich data, the interviewer must be able to develop a rapport with participants (262). Access to the BEAT-PCD network through European and international conferences facilitated the development of such rapport and underpinned the selection of interviewees, highlighting the appropriateness of this method for this study.

7.2.3 Study design and documents

We used a mixed-methods approach to evaluate differences between service delivery models. In order to do so, we adopted a two-stage design. The first step consisted of in-depth interviews with PCD specialists to provide a deeper understanding of the 'various realities' of service delivery for PCD. We were initially focused on European countries, in order to compare results with the 2009 survey. However, PCD experts from non-European countries demonstrated interest in participating in the study and we therefore decided to expand the scope of our research to include other countries.

Evidence of existing differences (or similarities) between models were gathered through interviews with PCD specialists from a variety of countries to uncover factors that might affect or determine variations in models of care. In the second stage, data gathered from the interviews informed the development of the survey, which was distributed internationally to PCD experts in order to quantify differences and similarities in delivery of care.

We developed two study-specific consent forms (one for the interview and one for the survey) and a participant information sheet. These were sent to volunteers prior to approaching them to participate in the study (Appendix E). Participants were asked to read these documents carefully and sign the consent form before the interviews were scheduled. The PCD specialists provided informed consent for digital recording of the interviews and were given the opportunity to review the full transcript, upon request. A separate consent was obtained before participants completed the survey.

7.2.4 Qualitative data analysis

Opportunist sampling is a common and acceptable approach in qualitative research in order to maximise resources available (263). Interviewees were derived from members of the BEAT-PCD network (194, 205, 206) and consisted of volunteers from different career stages, and invited experts from large PCD centres. The latter received a personal invitation by email to take part in the interviews, which were conducted in person, Skype or telephone calls.

Interviews were conducted by one person (BR) and lasted approximately one hour. We developed a semi-structured interview protocol, which served as a prompt for the interview to cover all aspects of service delivery (Appendix E E.1).

Interviews were recorded using a voice recorder and transcribed verbatim. Data were coded and analysed in its entirety using an inductive and semantic approach. The first line of coding described and summarised the data without interpretation of any underlying ideas. Data were

then organised in groups in the subsequent coding steps until common themes emerged. We used NVivo (version 11, QSR International, Doncaster, Victoria, Australia) for data analysis and applied the 6-phase guide advocated by Braun and Clarke (260) to conduct thematic analysis, using a constructionist approach.

7.2.5 Survey development and structure

The survey consisted of four parts, based on results from the qualitative analysis. The first one focused on characteristics of the respondents and included information on country, city, institution, title, speciality, age, sex, career stage, predominant duties, and involvement in research.

Part A consisted of repeated questions from the previous survey conducted in 2009. The selection of questions to be included in this section was based on discussions with collaborators that conducted the original survey. We made minor alterations to some of the categories in order to minimise the time needed to complete the survey. Part B were additional questions centred around service organisation and funding structure. The last section, part C, consisted of questions on the opinions of healthcare professionals involved in PCD delivery of care. These included opinions on barriers for patients to access the service, challenges in diagnostics and management of patients with PCD, areas that required improvements, and resources needed for the establishment of new PCD centres. The full survey is available on Appendix E E.2. Part A used the same wording from the previous survey, while questions and answers for parts B and C were based on the themes arising from the interviews.

Two healthcare professionals from UHS pilot tested the survey to ensure the questions and categories were understandable and relatable, and to record completion time. No adjustments were necessary as a result, and it took approximately 15 minutes to complete the survey.

7.2.6 Study population and quantitative data collection

We identified experts from PCD centres through the BEAT-PCD network. Paper-based surveys were distributed during the BEAT-PCD conference, which took place from 26th to 29th March 2019, in Poznan, Poland. Experts who were not present at the conference were invited through email to complete the online version of the survey, which was available at https://www.isurvey.soton.ac.uk/32325 from 2nd April to 6th June 2019. The link to the online survey was also circulated through the ERN network and leads of PCD centres were encouraged to forward it to colleagues in other countries that were not members of either networks. Non-respondents were sent a reminder email at the end of April 2019.

7.2.7 Choice of typology for country classification

To compare our results with those from the survey conducted in 2009, countries were grouped into five regions according to an adaption of the United Nations (UN) definition of European regions (257).

For the complete survey analysis, we conducted a comprehensive literature review on previously used country classification criteria in order to select the most suitable groupings, as the UN geographical criteria might not reflect the realities of service delivery for rare diseases. Several typologies have been developed throughout the years and the most relevant ones are detailed in Table 24 (Appendix E E.3). However, each classification has significant limitations; most importantly none included all countries that participated in the study (Table 24). We would therefore need to either exclude a considerable amount of countries or attempt to classify countries based on the selected criteria. Due to insufficient information provided by the studies and limited data available to the general public on each country's healthcare structure, the latter was not achievable. After careful review and discussions with collaborators, we opted to use an adaptation of the UN regional classification as it contained all countries included in the survey (257).

7.2.8 Quantitative data analysis

Study data were downloaded from isurvey as a comma-separated values (csv) file and uploaded to STATA for data coding, cleaning and recoding.

We obtained the complete cleaned and coded data and the data dictionary for the previous survey from our collaborators in Bern, Switzerland. We only included countries that were present in both surveys to allow for direct comparison of results. Data from Part A (Appendix E E.2) were coded to match the coding from the previous survey, therefore some of the variables had to be combined and recoded. For example, diagnostic testing data were captured as 'performed in centre', not in centre but 'sent to a national centre', 'sent to an international centre', or 'not performed' in the current survey. However, 'performed in centre' and 'sent to a national centre' were presented as a single category in the previous study; therefore, we merged the 'performed in centre' and 'sent to a national centre' categories in order to compare results from the two surveys. Additionally, two new variables were created by amalgamating two different variables: a) prophylactic nebulised antibiotics were merged with prophylactic oral antibiotics, and b) nebulised bronchodilators with metered dose inhalers were merged with those with powder devices.

We analysed and reported on aggregate data to keep the anonymity of PCD centre experts as they would be easily recognisable in countries where few PCD centres exist. Centres were initially grouped by country and then classified into geographical regions based on an adaptation of the UN defined regions (264). We merged Northern Africa and Western Asia to form the Africa & Eastern Mediterranean group, and Oceania with Eastern Asia and Southern Asia to form the Western Pacific & South Asia group. These changes were necessary as, using the original classification, some of the groups contained too few centres.

Additionally, we stratified PCD centres according to their size and the type of service provided (*i.e.* diagnostic or management) due to large variations in number of patients referred, diagnosed and managed at each centre, and based on findings from Strippoli *et al*, where size of centre was a significant co-variate. The complete survey data were then analysed with centres grouped by type and size (256).

Centres that diagnosed an average of less than five new cases per year were grouped as small diagnostic centres, those with five to nine new cases as medium-sized and those with at least 10 patients diagnosed per year as large diagnostic centres. When less than 30 patients were managed in the centre, these were classified as small management centres. Centres with 30 to 69 patients under care were considered medium-sized and those with at least 70 were large management centres.

Thresholds detailed above were selected based on Figure 22, with the aim of creating well-balanced groups (Table 13). Where we had more than one response per centre, we classified participants as either leads of diagnostic or management centre based on their reported predominant clinical duty. Therefore, we only included one response per centre when stratifying centres by size and specialty.

Categorical variables were reported as proportions. We used chi-squared test or Fisher's exact test where appropriate (*i.e.* when expected values in any cell of the contingency table was less than 5, or below 10 if only one degree of freedom) to compare categorical variables. P values <0.05 were considered statistically significant. Figures were plotted in Tableau 2019 v4.0. Missing data were not excluded from the analysis; they are shown in the figures where the bars do not reach 100%.

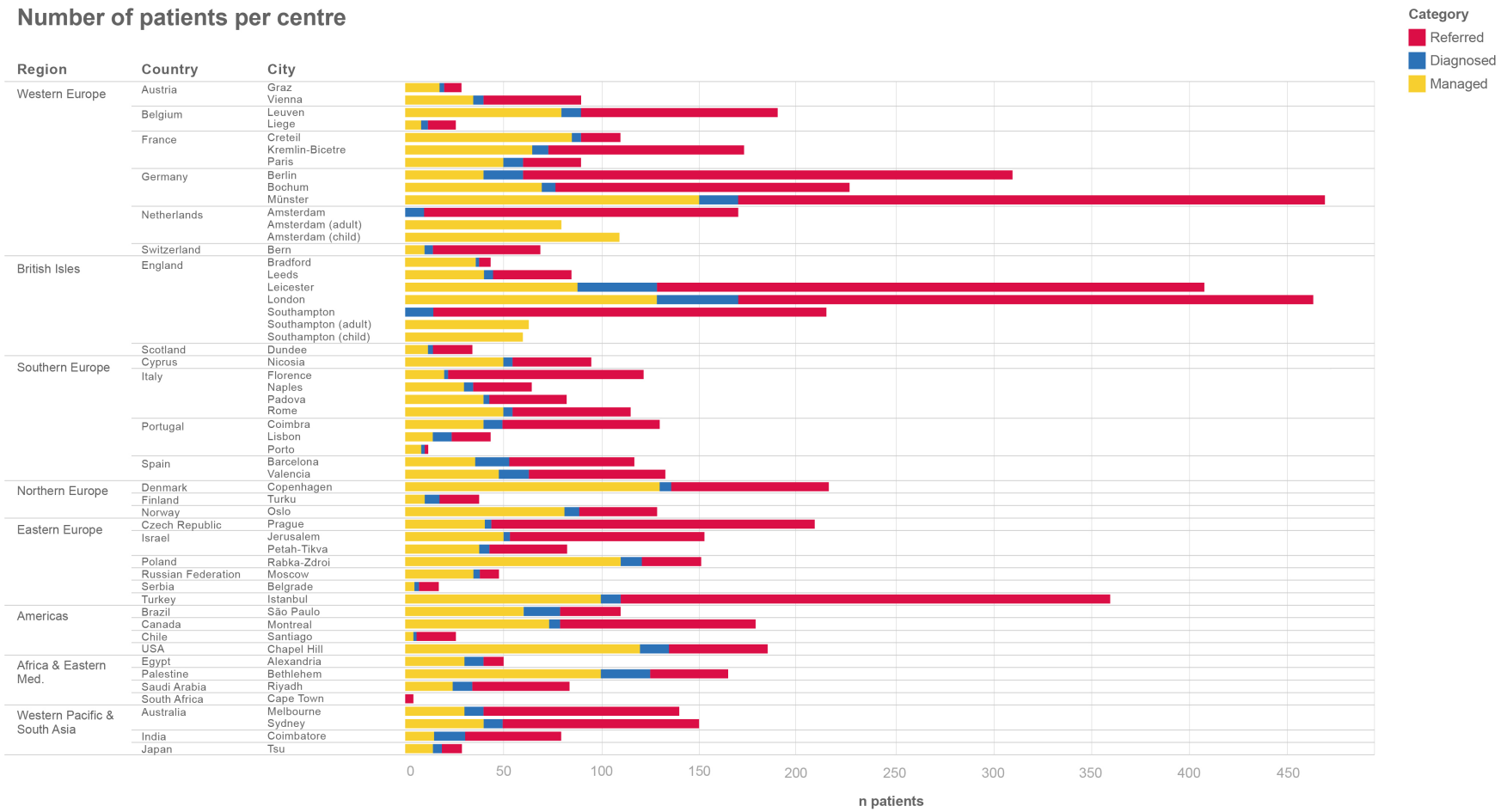


Figure 22. Average number of patients referred, diagnosed and managed at PCD centres per year, grouped by country and region.

7.3 Results

7.3.1 Qualitative research

Fifteen interviews were conducted between September 2016 and May 2018. Twelve self-volunteered and three were personally invited to participate in the study to expand on the geographical coverage and expertise provided by interviewees. Most participants were paediatric pulmonologists (*n*=9) and from European countries (*n*=10). Interviews lasted an average of 50 minutes (range 34 to 74 minutes) and were conducted face-to-face (*n*=5), telephone (*n*=2) or Skype (*n*=8).

The groupings highlighted in the figure above were used to develop the survey (see E.2).

7.3.2 Survey demographics

We invited 66 experts from PCD diagnostic and/or management centres in 32 countries. Fifty-eight completed the survey, a response rate of 87.9%. Greece was the only country without any respondents despite being invited to participate in the study. We excluded responses from two experts as those centres had more than one participant in the study. Where this happened, we selected the most senior participant, defined according to their career stage, to be included in the analyses.

Fifty-six experts from 31 countries were included. Most respondents were paediatric pulmonologists, independent of the size of the centre (Table 12). 61.5% of participants from small centres reported PCD diagnostics as their predominant duty, while most respondents from medium and big centres were involved in management of PCD patients as their predominant duty (Table 12).

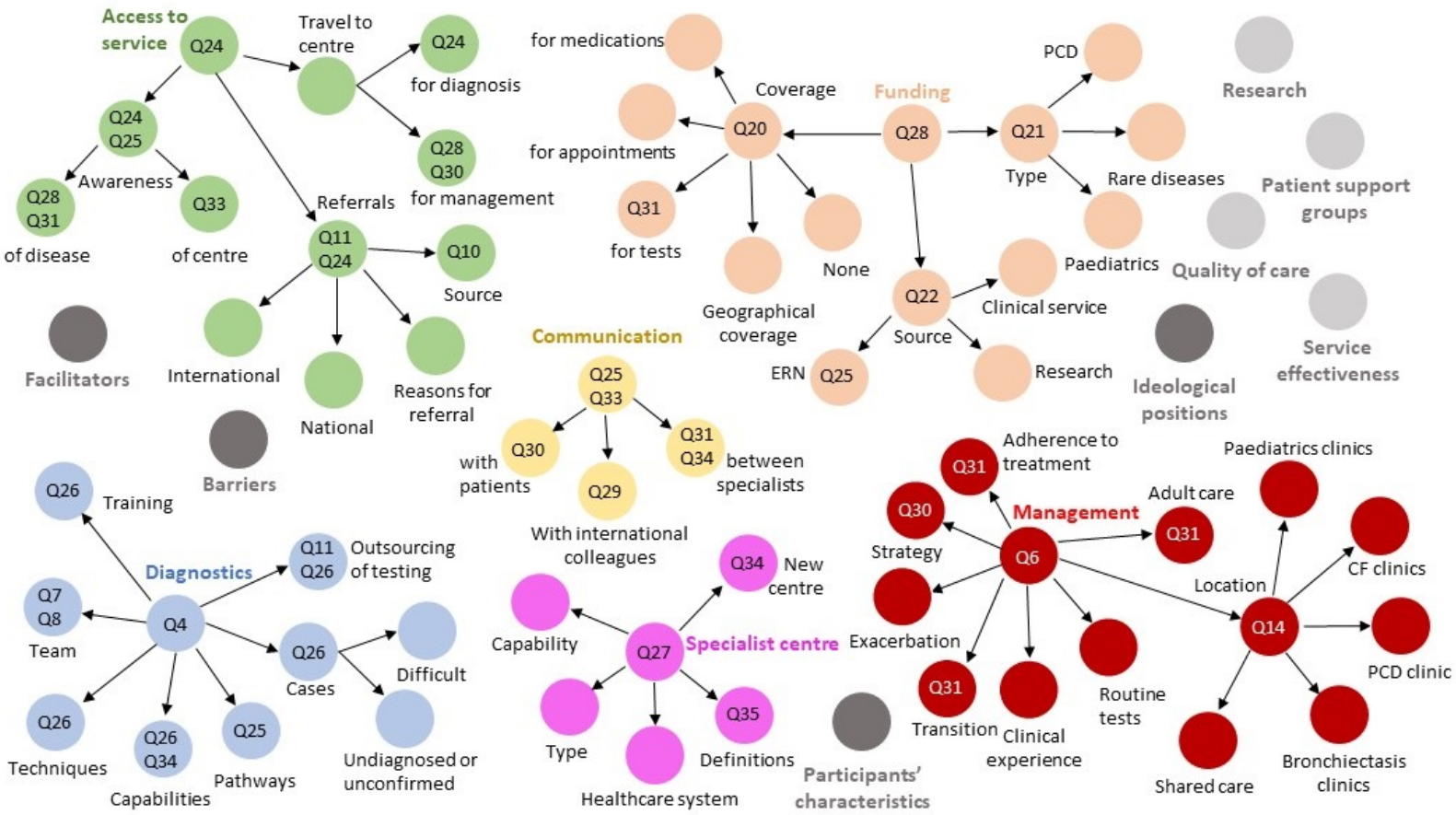


Figure 23. Mapping of themes derived from the interviews to questions added to the survey.. Each hub circle represents a theme, with arrows pointing to subthemes. Questions that were developed based on themes or subthemes are indicated inside each circle (e.g. Q4 inside the circle entitled 'diagnostics', in blue, indicates question 4 of the survey (see E.2)).

Table 12. Demographic characteristics of survey participants, stratified by size of PCD diagnostic centre

	Small centre ***n*** **(%)**	**Medium centre** ***n*** **(%)**	**Large centre** ***n*** **(%)**	**Total** ***n*** **(%)**
Specialty of respondents				
Paediatric	0	2 (9.1)	1 (6.2)	3 (6.0)
Paediatric pulmonology	10 (76.9)	11 (50.0)	12 (75.0)	33 (64.7)
Pulmonology	0	6 (27.3)	2 (12.5)	6 (12.0)
ENT	1 (7.7)	2 (9.1)	1 (6.2)	4 (8.0)
Fertility	1 (7.7)	0	0	1 (2.0)
Genetics	0	1 (4.5)	0	1 (2.0)
Diagnostic science	1 (7.7)	0	0	1 (2.0)
Career stage				
Early	0	1 (4.5)	1 (6.2)	2 (3.9)
Mid	3 (23.1)	5 (22.7)	0	8 (15.7)
Advanced	10 (76.9)	16 (72.7)	15 (93.7)	41 (80.4)
Age of respondents				
26-34 years	0	1 (4.5)	0	1 (2.0)
35-44 years	6 (46.2)	11 (50.0)	6 (37.5)	23 (45.1)
45-54 years	1 (7.7)	2 (9.1)	5 (31.2)	8 (15.7)
≥55 years	6 (46.1)	8 (36.4)	5 (31.2)	19 (37.2)
Predominant duty				
Diagnostics	8 (61.5)	5 (23.8)	2 (12.5)	15 (30.0)
Management	3 (23.1)	12 (57.1)	19 (62.5)	25 (50.0)
Laboratory researcher	1 (7.7)	1 (4.8)	0	2 (4.0)
Clinical researcher	0	3 (14.3)	4 (25.0)	7 (14.0)
Other	1 (7.7)	0	0	1 (2.0)

Most centres reported receiving patients from paediatric pulmonologists (72.7%) for diagnostic testing, followed by paediatricians (54.6%) and adult pulmonologists (34.6%). Centres reported a medium of 50 patients referred for diagnostic testing per year and seven new cases annually. The medium number of patients managed at PCD centres between 1st January 2017 and 31st December 2018 was 48 patients.

7.3.3 Complete survey analysis, stratified by geographical regions

Fifty-six experts from PCD centres in 31 countries completed the study survey (Figure 24). We received 14 responses from Western Europe, nine from the British Isles, 10 from Southern Europe, three from Northern Europe, seven from Eastern Europe, five from the Americas, four from Africa & Eastern Mediterranean, and four from Western Pacific & South Asia.

The diagnostic techniques available in each region, irrespective if preformed in centre or referred to a national centre for testing, are presented in Figure 25. Nasal NO was performed in most centres (85.8%). The exception was Africa & Eastern Mediterranean, where only one centre had nNO available.

Nasal brushing was the only technique conducted in all countries. Most centres in Europe went on to perform CBF (92.7%) and CBP (90.2%) by HSVA, and TEM (97.6%) from the nasal samples. HSVA was less performed in centres in the Americas (50%) and in Africa & Eastern Mediterranean, where CBF was not done in any of the centres and CBP was performed in two (50%) centres. On the other hand, TEM was available in 94.6% of centres, with no observed regional differences. Cell culture was rarely performed outside of the British Isles.

Immunofluorescence was available in all centres within the British Isles, in 78.6% of centres in Western Europe and half the centres in Western Pacific & South Asia. Almost all countries conducted genetic testing (96.4%), with no differences between regions. It was outsourced to another country in one centre in Southern Europe, one in the Americas and two in Africa & Eastern Mediterranean regions. All centres had either TEM or genetic testing available, in centre (80.4%) or through collaboration with other centres in the country (14.3%) or in other countries (5.4%).

Figure 24. Map highlighting the countries included in the survey.. Green circles indicate countries included in both the 2009 and 2019 surveys and therefore used for comparisons between surveys. Brown circles indicate additional countries included in the 2019 survey only.

Most centres (96.4%) encouraged patients to perform airway clearance techniques, following the ERS paediatric management consensus (265). Recommendation of formal exercise programme varied between regions, ranging from routinely recommended in 42.9% centres in Eastern Europe to none in Africa & Eastern Mediterranean, and from 60% in Southern Europe to 25% in Western Pacific & South Asia on a frequent basis (Figure 25).

All centres used antibiotics to treat exacerbations; however, it was rarely used when a positive culture was present in asymptomatic patients in 33.3% centres in Northern Europe, 23.1% in Western Europe, 20.0% in Southern Europe, and 11.1% in the British Isles. We found that regular intravenous antibiotics were never used in 41.5% centres and rarely in 45.3%, with significant variations between regions (p=0.016). On the other hand, intermittent antibiotics were used more frequently, ranging from 50.0% in Africa & Eastern Mediterranean to none in Southern Europe and in Western Pacific & South Asia.

We found that 92.6% of centres used nasal rinsing and 85.2% used hypertonic saline, either routinely or frequently. Hypertonic saline was used in all centres in the Americas and in Africa & Eastern Mediterranean. Bronchodilators were used routinely in 16.7% and frequently in 51.8% of centres, with all centres in Africa & Eastern Mediterranean prescribing it routinely, followed by two centres in Eastern Europe (28.6%) and one in the Americas (20.0%). Most centres (66.7%) rarely used inhaled corticosteroids. The most frequent prescribers were in Western Europe, Southern Europe, and Africa & Eastern Mediterranean, with one centre in each region using it routinely.

Tympanostomy with ventilation tube insertion was either rarely or never used in 85.2% of centres. It was mostly used in the Americas, where it was used either routinely (25%) or frequently (50%).

Most centres had follow-up visits with pulmonologists every three to four months (76.4%), independent of region. We found a wider variation in the frequency of follow-up appointments for physiotherapists and ENT physicians, with 53.7% and 9.3% respectively having regular visits every three to four months. Most ENT follow-ups were conducted annually (44.4%).

Leads from 17 countries reported having an established patient support group in their country, while 13 do not have a PCD-specific support group. One expert was not aware of the existence of patient support group in their country.

Figure 25. Diagnostic (top) and management (bottom) techniques stratified by geographical regions, described by participants of the study survey (2019).

7.3.4 Complete survey analysis by size of PCD centre

We found considerable variations between countries from the same region and between centres within the same country (Figure 22, Table 13). For example, the number of patients referred for diagnostic testing in Western Europe ranged from nine in Austria to 300 patients in Germany. In France alone, referrals varied from 20 in Creteil to 100 patients in Kremlin-Bicetre. The number of patients diagnosed in centres from the same geographical region also varied, with an average of only two patients per year diagnosed in Graz compared to 20 in Muenster. Large variations were also observed within regions for the number of patients managed at PCD centres. For example, in Western Europe 150 patients were managed in Muenster compared to only three patients in Liege. Due to these large variations within geographical regions, we opted to stratify centres according to size and type of service provided (see 7.2.8 for further details).

Large diagnostic centres were characterised by higher proportion of diagnostic tests available, in particular CBP by HSVA (85.7% vs 46.6% for medium and 6.7% for small centres), cell culture (47.6% vs 7.1% for medium and 6.7% for small centres), and immunofluorescence (42.9% vs none for medium and 20% for small centres). Medium-sized centres had the highest proportion of genetic testing available (71.4% vs 57.1% for large and 46.7% for small centres). However, both medium and large centres had genetic testing done within the country (92.9% and 89.5%, respectively). Genetic testing was outsourced to another country in one centre (7.7%) and was not available for two smaller centres (15.4%). Small diagnostic centres also had less access in centre to nNO measurements (66.7%) compared to medium (92.9%) and large centres (85.7%). Similarly, the proportion of small centres that performed TEM in centre was lower (40% versus 71.4% for medium and 76.2% for large centres, $p = 0.066$).

In terms of management strategy, large centres recommended formal exercise programme routinely (25%) or frequently (62.5%), while 23.1% of small management centres and 9.1% of medium-sized centres reported to never recommend it. All large management centres used nasal rinsing, either routinely (56.2%) or frequently (43.7%). Only 4.5% of medium-sized centres and 23.1% of small centre rarely recommended nasal rinsing (Figure 26). A third of centres reported using hypertonic saline routinely, more common in large management centres (43.7%) compared to medium (31.8%) and small centres (15.4%) (Figure 26). Inhaled corticosteroids were rarely prescribed in large centres (81.2%), while 46.1% of small centres used them frequently. Vaccine against *Influenzae* was recommended routinely or frequently in 87.5% of larger centres and 81.8% in medium centres, compared to 76.9% in smaller centres. None of the management centres used

rhDNase routinely, with only one centre in Southern Europe and one Eastern Europe reporting frequent use.

Prophylactic antibiotics were prescribed more frequently in large management centres (68.7% compared to 54.5% for medium and 38.5% for small, Figure 26), as was intermittent intravenous antibiotics (6.2% routinely and 43.7% frequently in large centres, 4.8% routinely and 19% frequently in medium centres, and none routinely and 23.1% frequently in small management centres). Use of regular intravenous antibiotics was more common in large centres, where they were prescribed frequently in 18.7% of centres compared to 9.5% in medium centres and none in small centres.

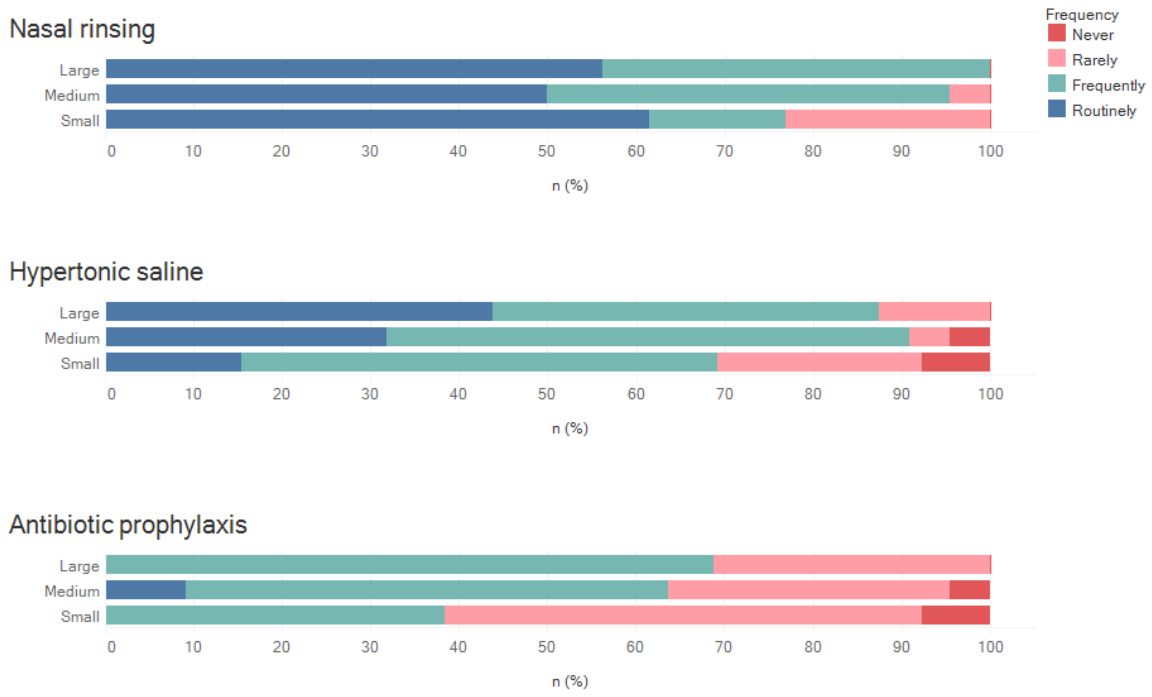


Figure 26. Use of nasal rinsing, hypertonic saline and antibiotic prophylaxis across management centres, stratified by size of management centre.

Most participants described their country's PCD service as centralised, particularly for diagnostics (80% vs 66.1% for management services), with no significant difference in size of centre and type of service provided ($p = 1.00$ for diagnostic centres and $p = 0.23$ for management centres).

We found that 57.1% of respondents identified their healthcare system as public, followed by a combination of public and private (mostly public with some private in 30.4% of cases). Only one centre selected the private option. There were no significant differences when stratified by size of diagnostic or management centres ($p = 0.58$ and $p = 0.38$, respectively). However, the cost of the appointments for most patients seen at PCD centres were covered by public funding (91.1%), through taxes or compulsory social insurance.

Large and medium-sized diagnostic centres received national funding for PCD services (66.7% and 57.1%, respectively). In contrast, only four of the small diagnostic centres (26.7%) received national funding for PCD services. A quarter of centres received additional research funding, more common in small and large diagnostic centres (41.7% each compared to 16.7% in medium-sized centres).

Half of the diagnostic centres (51.0%) provided free treatment to PCD patients, irrespective of being assigned a diagnosis of PCD. Small centres received additional funding from the government if a diagnosis of PCD was confirmed (40.0%), while medium-sized centres received additional provisions if either a PCD or rare disease diagnosis was confirmed (23.1% each).

7.3.5 Comparison with previous survey

For comparisons between the two surveys, we excluded data from eight countries (*i.e.* Ireland, Greece, Sweden, Bulgaria, Estonia, Hungary, Romania and Slovakia) from the previous survey, and responses from 13 from the current survey, including two European countries that were not included in the 2009 study (*i.e.* Poland and Russia). We included responses from 18 countries (131 participants in the previous survey and 41 participants in the survey conducted in 2019) (Figure 24). Despite the difference in sample sizes, the proportion of centres contributing data from each UN region were similar (Table 14).

Table 13. Distribution of centres that completed each survey, stratified by UN European region

	Previous survey (2009) (256)	**Current survey (2019)**
Western Europe	45 (34.3%)	14 (34.1%)
British Isles	18 (13.7%)	9 (21.9%)
Southern Europe	37 (28.2%)	10 (24.4%)
Northern Europe	8 (6.1%)	3 (7.3%)
Eastern Europe	23 (17.6%)	5 (12.2%)
Total	131 (100%)	41 (100%)

Table 14. Average number (%) of patients referred, diagnosed and managed at PCD centres per year, stratified by region

	Small diagnostic centre n (%)	**Medium diagnostic centre n (%)**	**Large diagnostic centre n (%)**	**Total diagnostic n (%)**	**Small management centre n (%)**	**Medium management centre n (%)**	**Large management centre n (%)**	**Total mangement n (%)**
Western Europe	3 (25.0)	4 (33.3)	5 (41.7)	12 (100)	3 (23.1)	4 (30.8)	6 (46.15)	13 (100)
British Isles	2 (33.3)	1 (16.7)	3 (50.0)	6 (100)	1 (14.3)	4 (57.1)	2 (28.57)	7 (100)
Southern Europe	3 (30.0)	3 (30.0)	4 (40.0)	10 (100)	3 (30.0)	6 (60.0)	1 (10.00)	10 (100)
Northern Europe	0	3 (100)	0	3 (100)	1 (33.3)	1 (33.3)	1 (33.33)	3 (100)
Eastern Europe	4 (57.1)	1 (14.3)	2 (28.6)	7 (100)	1 (14.3)	4 (57.1)	2 (28.57)	7 (100)
Americas	2 (50.0)	1 (25.0)	1 (25.0)	4 (100)	1 (25.0)	1 (25.0)	2 (50.00)	4 (100)

	Small diagnostic centre n (%)	**Medium diagnostic centre n (%)**	**Large diagnostic centre n (%)**	**Total diagnostic n (%)**	**Small management centre n (%)**	**Medium management centre n (%)**	**Large management centre n (%)**	**Total mangement n (%)**
Africa & Eastern Med.	1 (25.0)	0	3 (75.0)	4 (100)	1 (33.3)	1 (33.3)	1 (33.33)	3 (100)
Western Pacific & South Asia	0	1 (25.0)	3 (75.0)	4 (100)	2 (50.0)	2 (50.0)	0	4 (100)
Total	15 (30.0)	14 (28.0)	21 (42.0)	50 (100)	13 (25.5)	22 (43.1)	16 (31.37)	51 (100)

The number of diagnostic referrals per year increased from a median of 12 in 2009 to 60 in 2019 across all countries; however, the number of new cases diagnosed annually decreased from 18 to six. Reasons for the decline in the number of new cases diagnosed per year are unclear. We found a global increase in the number of centres that performed nNO testing, nasal brushing, TEM and genetic testing (Figure 27).

All centres in Western Europe, British Isles and Northern Europe performed nNO testing, however three centres in Southern Europe and one in Eastern Europe still did not measure nNO levels. The proportion of centres performing cell culture in Western Europe, British Isles, and Southern Europe increased, however it decreased in Northern and Eastern Europe. All centres were performing nasal brushing and virtually all had access to TEM and genetic testing in centre, a significant increase from ten years ago ($p < 0.001$ for both diagnostic tests). The number of centres in Western Europe, British Isles and Southern Europe performing CBF and CBP by HSVA increased, while outsourcing of diagnostic tests decreased, with only cell cultures in Eastern Europe and genetic testing in Southern Europe currently performed in another country.

The number of patients managed under each centre increased from a median of four to 50 patients. The most striking difference between the two surveys was the significant increase in the use of hypertonic saline ($p<0.001$), seen in all European regions. We found an increase on the routine or frequent use of bronchodilators in the British Isles, likely associated to the concomitant use of hypertonic saline. There was a notable decrease in the frequent use of inhaled corticosteroids in all regions except the British Isles and Eastern Europe (Figure 28).

We observed an increase in the frequent use of prophylactic antibiotics in all regions, more pronounced in Western Europe and the British Isles. One centre in Northern Europe used regular intravenous antibiotics routinely, while centres in other regions rarely or never used them. There was no change in the use of antibiotics to treat exacerbations compared to the previous survey, with most centres continuing to use it routinely.

We found a global decrease in the frequent and routine insertion of tympanic membrane ventilation tubes. However, it was still used as treatment to recurrent chronic otitis media in all regions except Northern Europe, where tympanostomy was not performed.

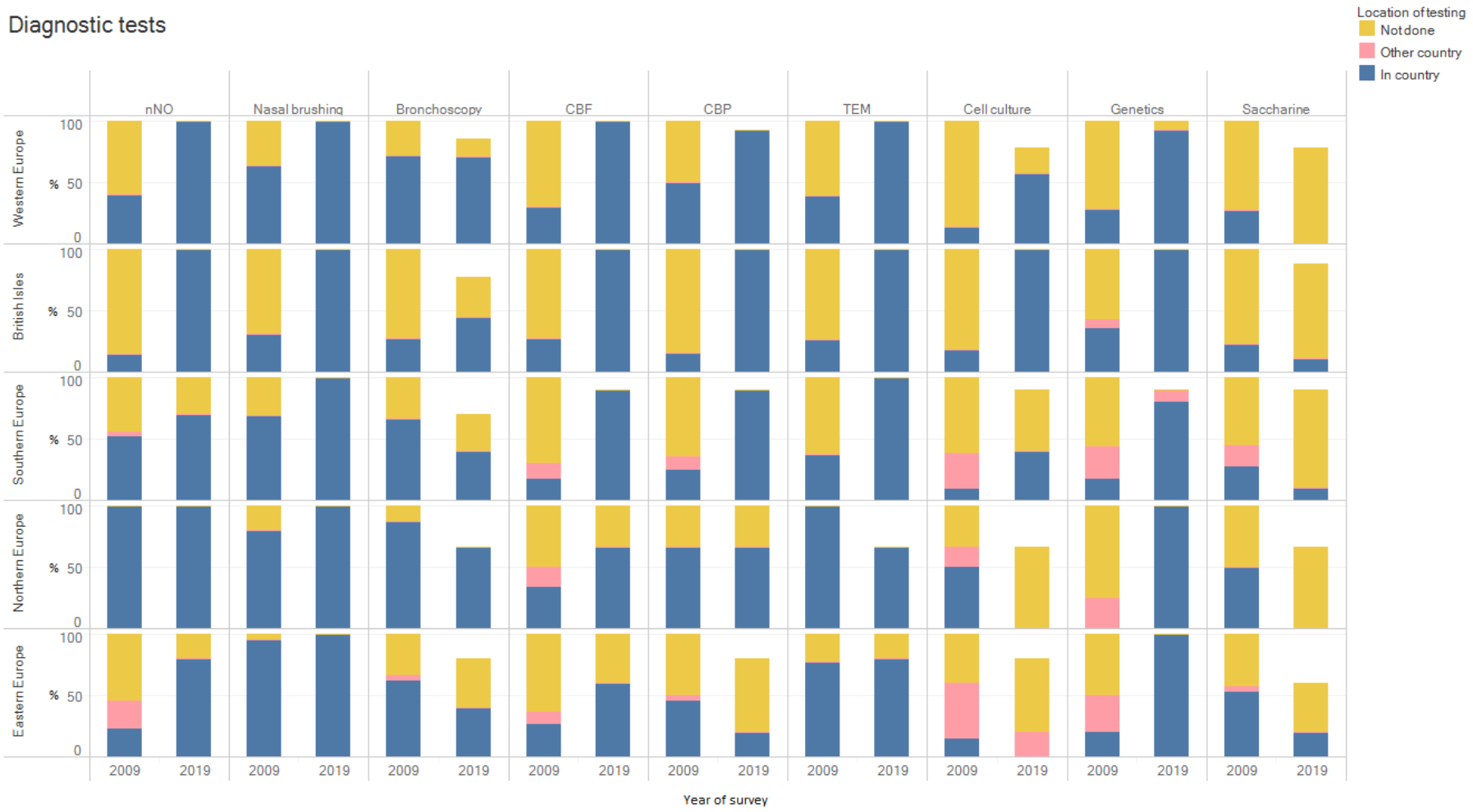


Figure 27. Comparison of diagnostic tests available according to the 2009 and 2019 surveys, stratified by European region.

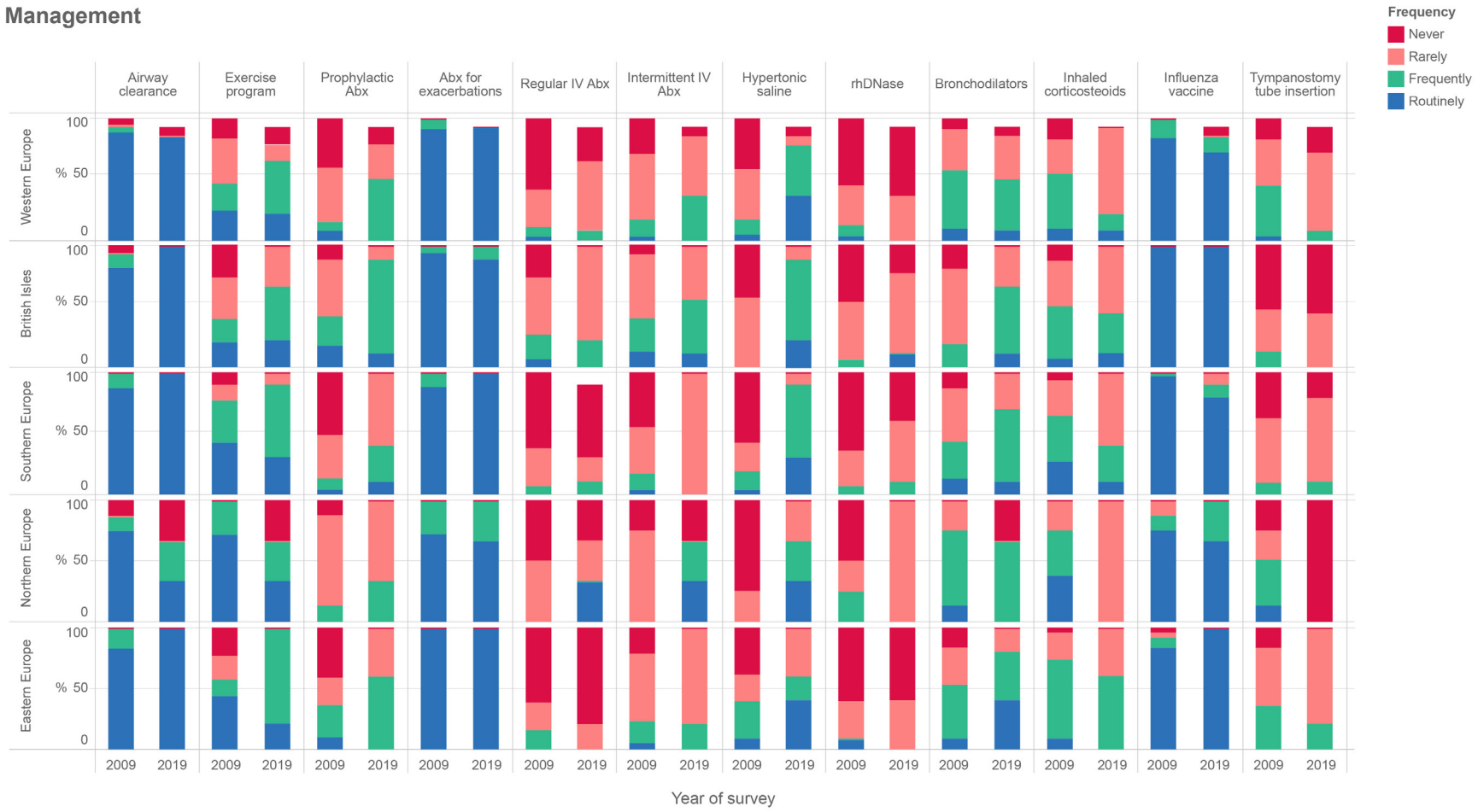


Figure 28. Comparison of management strategies according to the 2009 and 2019 surveys, stratified by European region.

7.4 Discussion

7.4.1 Diagnostic strategies

We found small variations between regions in availability of diagnostic tests, particularly between European countries. Most centres measured nNO levels, however it was generally not available in centres in Africa & Eastern Mediterranean region. HSVA was commonly performed in European centres, while only half of the centres in the Americas reported access to the test in centre. In North America it has a less prominent role in PCD diagnosis, with the ATS guidelines advocating against its use (65). This reflects intrinsic differences in the healthcare structure of these countries. For example, insurance companies in the US demand processing of samples in centre, which is challenging for HSVA as it requires experienced personnel and expensive equipment.

There was a global increase on the availability of diagnostic tests in PCD centres compared to the survey conducted in 2009, which showed that 16% of centres did not conduct ciliary function tests or TEM (256). Our study showed that the two confirmatory tests according to the ERS and the ATS diagnostic guidelines, TEM and genetic testing, were widely available, indicating that PCD diagnosis has become more standardised in the last decade (25, 65). All but three centres reported having access to TEM, either in centre, in country or through collaboration with other countries, and there was a sharp increase in the proportion of centres that had access to genetic testing. Immunofluorescence, a relatively new techniques which was not captured in the 2009 survey, was now available in over 40% of large diagnostic centres and in all centres in the British Isles (88).

The decrease observed on diagnostic tests that are outsourced to other countries across Europe reflects the restructuring of PCD diagnostic services in the last decade. The clearest example is England, where diagnostic services are now centralised around three nationally funded diagnostic services (see 3.2) (3, 6). Resources and availability of diagnostic tests are now concentrated around large diagnostic centres, independent of geographical region.

7.4.2 Management strategies

Most centres used airway clearance techniques and encouraged PCD patients to practice physical activities, following the strong recommendations made by the 2009 ERS paediatric management consensus (102). The increase in the use of hypertonic saline in the last ten years likely reflects both the fact that it is cheap and widely available, and the current evidence indicating a perceived improvement on HRQoL compared to placebo (106), as discussed in 6.3.9. The expansion of use of

bronchodilators in the British Isles was probably associated with the adjacent use of hypertonic saline, which has also increased, as bronchodilators are commonly used prior to nebulised hypertonic saline to mitigate bronchoconstriction responses. The ERS consensus guidelines recommended restricting the use of bronchodilators to cases where airway reversibility can be seen on spirometry (102). As discussed in 6.3.9, crossover trials found no evidence that its use improves lung function or delays disease progression.

There was a wider variability of management strategies when stratified by size of centre, likely due to the lack of evidence-based guidelines. All large management centres used nasal rinsing routinely or frequently, and hypertonic saline was prescribed more often than in medium-sized and small management centres, treatments often considered beneficial in PCD. On the other hand, almost half of the small management centres were still prescribing inhaled corticosteroids frequently despite lack of clinical evidence and little theoretical benefit unless the patient also has asthma. We did not observe any differences on the use of inhaled corticosteroids by geographical region. Since leaders in the field are generally concentrated in larger specialist centres, these findings suggest that clinical experience managing PCD patients is a major factor in choice of management strategy.

Recently published findings from a large international RCT found that the use of prophylactic Azithromycin significantly reduced the number of pulmonary exacerbations (266). We found a global increase in the use of prophylactic antibiotics in European countries compared to 10 years ago, particularly in Western Europe and the British Isles. Large management centres used prophylactic antibiotics, and intermittent and regular intravenous antibiotics more frequently than smaller centres. Despite the unclear role of the use of antibiotics when a positive culture is found in patients without any clinical indicators, it was extensively used routinely or frequently throughout PCD centres, independent of geographical region or size of management centre.

We observed a decrease in the frequency of routine or frequent tympanostomy in European centres compared to the 2009 survey, likely as a result of the ERS consensus advising against the insertion of ventilation tubes in children due to the increased risk of the development of offensive otorrhea (102). In contrast, centres in the Americas routinely or frequently opted for the surgical treatment, as the North American consensus does not recommend against the use of ventilation tubes but rather promotes counselling of patients on the potential risk of prolonged otorrhoea post-ventilation tube insertion (103).

7.4.3 Referrals and funding structure

The increase on cases referred to PCD centres for testing strongly suggests that awareness of the disease by referring physicians, particularly paediatric pulmonologists, has improved. This was also shown by Rubbo *et al* (6), where data from 333 children undergoing annual review in 2015 at one of the four PCD paediatric management services in England showed that age at diagnosis was 2.6 years, considerably younger than previously reported in the literature (4). It also suggests that there has been a shift from decentralised to centralised PCD services, reflected by the steep increase on number of patients managed under each centre compared to ten years ago. While only three countries had centralised services for PCD in 2009, most participants in the current survey described the service provided in their country as centralised, particularly for diagnostic testing. The need for highly specialised and expensive equipment, and trained diagnostic scientists and technicians underscores the advantages of concentrating resources in a selected number of centres per country (93), as discussed in 2.1.1.

Funding structures and healthcare systems are important contextual evidence and can partially explain the heterogeneity of health services. Over half of large and medium-sized centres received national funding, while only approximately a quarter of small centres had financial support from central government. Additional provisions were available when PCD diagnosis was confirmed in smaller centres, whereas medium centres seemed to be structured around confirming a diagnosis of rare disease. Half of the centres included in the current survey were in countries with a public healthcare system. One participant identified their country's healthcare system as private, in which case the choice and availability of diagnostic testing might be influenced by insurance coverage. For example, genetic testing is considered a reliable and regulated test in North America and therefore recognised by insurance systems as a PCD diagnostic test (92).

Additionally, geographical spread of specialised centres can impact on the accessibility, availability and feasibility of testing, and these are mirrored by the differences in recommendations made by the European and North American diagnostic guidelines (see 2.4.6). Structural differences are particularly relevant to inform the planning of new PCD centres, as healthcare system and funding pathways usually cannot be altered.

7.4.4 Potential implications

Data from this study can be used by healthcare professionals and policy makers at all levels of government to a) inform the establishment and development of new diagnostic and/or management centres for PCD, and b) improve the delivery of care for patients accessing existing

PCD centres. It can also be used to set a wider agenda for rare lung diseases, through European collaboration (*e.g.* ERN-LUNG).

Importantly, this study represents the initial step in describing differences in service delivery models across countries. Our findings will hopefully foment the undertaking of new studies in this vastly under researched area.

7.4.5 Strengths and limitations

This study was the first to describe several aspects of service delivery across multiple PCD centres. We grouped countries both by geographical region and centre size. The former was necessary to compare results with the survey conducted a decade ago and the latter was a consequence of the wide variations observed between neighbouring countries and within countries. We had a unusually high response rate, likely due to the establishment of highly collaborative PCD networks and the data collection strategy, which included distributing the survey to experts during the annual PCD conference. Another strength of this study was the availability of coded data for re-analysis from the 2009 survey. This provided the opportunity to directly compare results from 2009 to the findings obtained from this survey.

Our study had significant limitations. We used different data collection strategies to the 2009 survey due to advances on delivery of care for PCD in the last decade. Participants were approached in person or invited to complete an online survey, while the previous survey relied on mailed paper-based questionnaires. These differences may have introduced selection bias, as respondents to this study's survey were more likely to be involved in PCD delivery of care and research on a daily basis. However, we had to adapt our data collection strategy to reflect the advent of PCD networks and the centralisation of services in highly specialised centres.

The small sample size limited further comparisons with the 2009 survey, a result of the centralisation of services that has occurred in the last decade. Additionally, due to our approach targeting members of the BEAT-PCD network, we had a higher proportion of European centres. This might reflect the current state of PCD delivery of care, with patients in countries outside of Europe, North America and Australia having limited access to diagnostic testing and, therefore, many remain undiagnosed. Another limitation was the possibility that countries with decentralised services were underrepresented in our study, as non-European countries that were not members of any of the established networks might have not been aware of the survey. We attempted to reach experts in centres that are not active within the BEAT-PCD network by circulating the survey invitation and link through the ERN mailing list and requesting that participants invite other leads in their country, where appropriate.

This study captures the availability of diagnostic tests and the choice of management strategy as viewed by PCD specialists so it might not represent true clinical practice. Additionally, questions and categories were based on interview data from a selected number of PCD specialists and their account through interviews might not necessarily represent a direct expression of their experiences. In fact, interviews often represent the individual's account through their own interpretation and expectations of what they believe will be viewed in a positive manner by the interviewer. The mere presence of a formal process in which participants are asked to speak about a specific topic is likely to influence responses given, particularly when there is a professional relationship between the parties involved. However, this limitation would not be overcome by adopting any other data collection method, or by simply mirroring the UK context, which would not be applicable to centres in other countries.

We collected qualitative and quantitative data on the opinion of PCD specialists on service delivery for PCD; these will be analysed as a separate project and will provide important contextual and explanatory evidence to elucidate some of the variations reported by this study.

7.5 Conclusions

Diagnostic testing for PCD has become more standardised and centralised in the last decade due to the advent and implementation of evidence-based diagnostic guidelines (25). However, there is a clear need for evidence-based PCD management guidelines. Variations in choice of management strategy were more evident when centres were stratified by size of centre. Small management centres continued to use therapies that are commonly used in other diseases but have little theoretical and no apparent benefit in PCD, such as bronchodilators and inhaled corticosteroids. Larger centres were more likely to routinely or frequently use simple techniques that are considered effective to manage PCD patients such as airway clearance, formal exercise programme, nasal rinsing and immunization.

This study provided a comprehensive review of the current state of service delivery for PCD. Our data can be used to inform healthcare professionals and policy makers at the European, national and local levels to make informed decisions on the best model for their type of service based on evidence, as opposed to subjective views and personal preferences. Further research is needed in order to better understand how different models for delivery of care can impact on patient satisfaction and on clinical outcomes. They can also inform how different models can be adapted to countries with limited resources and countries setting up new PCD centres.

Chapter 8 **General discussion**

In this thesis, I explored different aspects of delivery of care for PCD. Figure 29 summarises the main themes that have explored throughout the thesis; these will be further discussed in the next sections. Additional challenges, common to other rare diseases, have also been explored throughout the chapters, such as underdiagnosis, wide geographical distribution of cases, absence of core disease-specific outcome measures, variations on management strategies, lack of disease-specific treatments, and small sample sizes requiring international collaboration and novel methods to conduct meaningful studies.

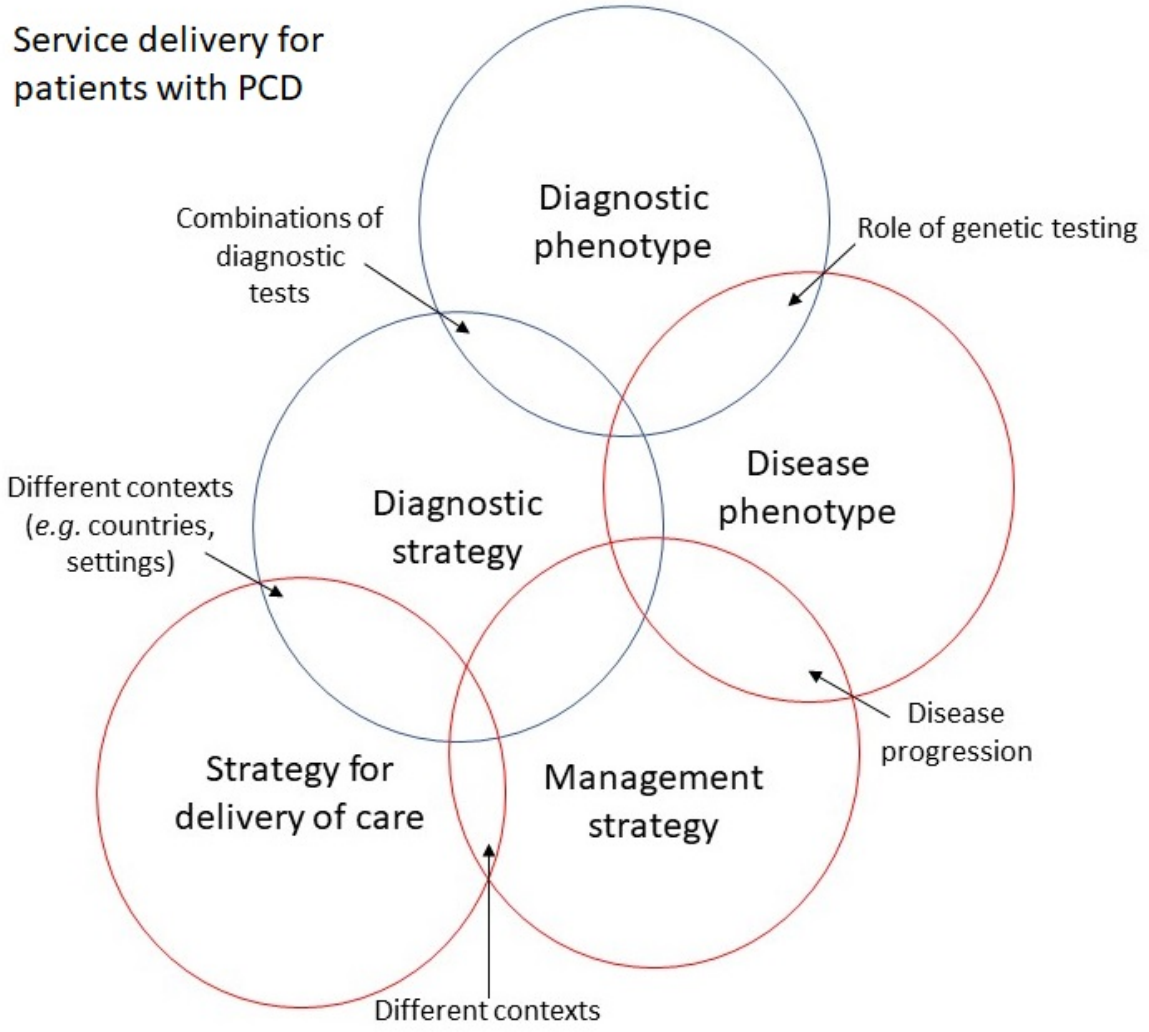


Figure 29. Venn diagram summarising the service delivery framework adopted in this thesis.. Circles depicted in blue represent the themes under the diagnostic sphere, while those in red represent management themes.

8.1 Diagnostic phenotype

In Chapter 4 we showed that HSVA, the only clinically available diagnostic test that directly assesses ciliary function, is an accurate and reliable test when done by experienced scientists. Despite being a genetic disease, there is no single gold standard test to diagnose PCD as many patients do not have known mutations in PCD-related genes. The current ERS diagnostic guidelines acknowledges the important role of HSVA in diagnostic testing but recommended at least three brushings or two brushings followed by cell culture to confer a 'PCD highly likely' diagnosis.

According to evidence-based guidelines, diagnosis should be based on combination testing, with hallmark ultrastructure seen on TEM and bi-allelic mutation in known causative gene the only two confirmatory tests to date. However, the sensitivity and specificity of these diagnostic tests needs to be considered. TEM was found to have a sensitivity of approximately 0.7 and specificity of 0.9, while genetic testing can identify up to 70% of cases. Therefore, both tests miss a considerable number of positive patients. Additionally, interpretation of genetic testing results depends on several factors, including number of genes tested, testing method or technique, and standardised reporting of results. Similarly, TEM findings rely on adequate sample collection, processing and examination by experienced scientists. Chapter 4 provides new evidence to strengthen the role of HSVA in the diagnostic pathway.

Mutations in genes that encode ciliary proteins result in ultrastructural defects that can usually be visualised by TEM or electron microscopy tomography and produce ciliary beat pattern abnormalities seen by HSVA. The TDA models depicted in Chapter 5 showed that HSVA and TEM results can guide genetic testing as the gene groups mapped closely to CBP by HSVA and ultrastructural defect by TEM. Levels of nNO were discriminatorily lower in patients with mutations in genes involved in dynein structure compared to X-linked genes and those involved in ciliogenesis.

Additionally, a significant number of variants of unknown significance and single heterozygous mutations can be found in patients with a suspected clinical phenotype and therefore TEM and HSVA can provide additional information to confirm cases that are unresolved by genetic testing. Findings from Blanchon *et al* (178) corroborate our conclusions, with the authors advocating use of HSVA to guide diagnostic testing. Results from other diagnostic tests can also assist in providing the pathogenicity classification, particularly for novel variants.

8.2 Diagnostic strategy

Findings from Chapter 4 should be interpreted in light of the service delivery model adopted in each country, including contextual evidence such as healthcare system and availability of resources, as centres have different diagnostic settings and expertise. Kouis *et al* (267) investigated varied combinations of sequential and parallel testing to determine the most effective and less costly combination of tests. This is particularly important for countries with low resources, where allocation and coordination of available diagnostic tests is crucial (93). Chapter 7 showed that small diagnostic centres had less access to nNO, TEM and genetic testing, while all tests were available in large diagnostic centres, including cell culture and immunofluorescence.

Improvement of diagnostic testing relies on effective referral pathways to correctly identify cases with a high likelihood of having PCD. These should be referred for testing in centres with appropriate equipment and expertise. Disease awareness seems to be improving, as shown in Chapter 7, where our survey found a median of 60 referrals per year compared to 12 reported by the survey conducted in 2009. This increase in referrals might be a consequence of the centralisation of PCD diagnostic services, with 80% of survey participants identifying their diagnostic service as centralised. The development of evidence-based guidelines with recommendations on criteria for patient referral and the use of symptoms-led clinical scores likely had a positive impact on the referral pathway (17, 57).

Our findings in Chapter 4 allow for early initiation of treatment and counselling of patients that have abnormal ciliary function seen on HSVA on the day of their appointment, decreasing the diagnostic delay. Results from confirmatory tests, on the other hand, can take up to six months to become available. In Chapter 5 we showed that mutations in certain genes were associated with worse respiratory phenotype at diagnosis. Since early diagnosis can have a positive effect on delaying disease progression, there might be an added benefit to move genetic testing up the diagnostic pathway in order to identify genes that are correlated with worse clinical outcomes. It might also reduce the diagnostic delay in certain contexts, such as countries without easy access to TEM and HSVA or with decentralised diagnostic services. However, it is important to stress that interpretation of genetic results requires geneticists experienced in PCD and should be interpreted in light of results from other diagnostic tests. The role of genetic testing in PCD, therefore, is still limited and should not be conducted in isolation but in tandem with other diagnostic tests.

8.3 Disease phenotype

Recent studies have focused on the effect of different genotypes on disease phenotype. In Chapter 5 we confirmed the findings of Davis *et al* (28) by demonstrating that patients with mutations in *CCDC39* had lower FEV_1 z-scores at diagnosis. Furthermore, we showed that these patients also had a significantly higher proportion of history of NRDS compared to those with other mutations, suggesting that respiratory symptoms can be seen from birth and progress into lung impairment at an early age. However, cohort studies are required to determine if these findings can be validated in a longitudinal model. Subsequent findings by Davis *et al* (180) suggest this is the case but, like in the original study, patients were grouped by TEM findings as opposed to actual genotype, limiting conclusions.

In Chapter 5 we also demonstrated that patients with *DNAH11* mutations had significantly higher FEV_1 z-scores at diagnosis and less NRDS compared to patients with other mutations. These are novel findings and might inform frequency of appointment and aggressiveness of treatment for patients with these mutations. In contrast, patients with mutations in the most common genetic cause of PCD, *DNAH5*, had no distinctive phenotypic characteristic compared to patients with mutations in other genes. However, the use of FEV_1 z-scores might not be appropriate to detect early lung damage as it reflects proximal airways impairment, as highlighted in Chapter 6.

Previous studies have shown that FEV_1 decline is heterogenous, with a steeper rate seen in patients with chronic *Pseudomonas aeruginosa* infection (122). However, variations in the definition of chronic colonisation discussed in Chapter 6 are common in the PCD literature and can affect findings. Efforts to standardise these terms are urgently needed. Additionally, measurement of spirometry-derived parameters should follow standardised protocols, with findings from Chapter 6 underscoring that only approximately 44% of the studies that used FEV_1 as an outcome measure reported applying ATS/ERS standards to perform spirometry measurements (123).

Some of the findings suggested by the discovery model depicted in Chapter 5 were not validated in the subsequent topological model. For example, we observed a clear inverse association between having lower airway symptoms and history of upper airway symptoms in the discovery model. Additionally, two distinct clusters were seen: a group of patients with mutations in N-DRC and molecular ruler genes that did not have a history of rhinitis, and patients with mutations in genes related with the structure and assembly of the dynein arms that did not present with a history of glue ear. There are several possible explanations for the fact that these findings were not replicated. We used a standardised data collection form with definitions of variables for the recording of phenotypic characteristics but differences in availability of information and definition

of commonly used clinical terms might explain some of the variations between the findings from the discovery and validation topological models. There were differences in the population included in each dataset, as the validation model included patients from France and the Netherlands in additional to those from the UK. Alternatively, non-replicated findings from the discovery model might be due to chance; however, they should be investigated in classic hypothesis-testing studies.

8.4 Management strategy

In line with results of Chapter 5, management strategy will likely be guided by genotype-phenotype associations in the future. However, current management often relies on evidence derived from CF and non-CF bronchiectasis studies as there is no PCD-specific management evidence-based guidelines to date.

It is widely accepted that respiratory management of PCD relies on promoting airway clearance, impeding or delaying lung disease progression, and treatment of acute pulmonary infections. However, in Chapter 7 we show that there are variations on management strategies. Centres that manage less than 30 patients were less likely to prescribe airway clearance techniques and formal exercise programmes compared to centres that manage at least 70 patients. In fact, smaller management centres prioritise treatment options that are effective for CF or non-CF bronchiectasis patients such as inhaled corticosteroids and beta agonists but were found to be ineffective in PCD.

Monitoring of lung function is often done through spirometry measurements. In Chapter 6, our scoping review showed that spirometry-derived measures were the outcomes most often used, highlighting the availability of these data for longitudinal monitoring. Differences in standardisation on reporting of spirometry measurements were considerable, with only a fifth of the studies using FEV_1 z-scores, which are centralised and adjusted for height, sex, age, and ethnicity.

Low-radiation modules of chest HRCT might be a helpful tool for longitudinal assessment of PCD patients, as it can detect structural changes early in life. It has been proposed as a surrogate marker for lung disease but there are no PCD-specific scoring systems to date, with studies using modified CF scores. In Chapter 6, our attempts to compare studies that used HRCT to other tests were hampered by the lack of standardised measurement and reporting of HRCT-derived parameters. This is a significant setback as PCD research moves towards RCTs and prospective cohorts, highlighting the need for a core set of outcome measures that are consistently measures across different studies.

Equally important is to use clear definitions for commonly applied terms. For example, variations in the definition of chronic colonisation, as discussed in Chapter 6, are common in the PCD literature. Attempts to develop disease-specific definitions such as the exacerbation consensus are essential to generate comparable findings.

The use of antibiotics to treat exacerbations is recommended by both the ERS and the ATS consensus statements. In Chapter 7, we showed that almost all PCD centres follow these recommendations, irrespective of geographical region or size of management centre. Recommendation on the use of antibiotic prophylaxis are less clear and variation on its use were more pronounced, with small management centres less likely to prescribe it on a frequent basis compared to medium and large centres. Most PCD centres prioritise techniques that are cheap, widely available and safe such as nasal rinsing and hypertonic saline, despite lack of conclusive evidence of its benefit in PCD.

8.5 Strategy for delivery of care

Improvements in delivery of care for rare diseases has the potential to impact outcomes across different centres and enhance the quality of life of affected individuals. We showed in Chapter 4 that HSVA is an accurate and reliable test to diagnose PCD when done by scientists with over ten years of experience in ciliopathies. Due to small number of patients, limited knowledge on disease-specific diagnostic and management techniques by generalists, and lack of resources in many countries, it would be unrealistic to expect that all secondary and tertiary care centres provide the best possible healthcare for patients with rare diseases. There has been a shift from decentralised systems of delivery of care for PCD to centralised models in order to concentrate resources such as specialised equipment and experienced staff in a few centres, as shown in Chapter 7. The same has occurred in CF, with the creation of CF Care Centres in the USA centralising and standardising the delivery of care for patients with this rare condition (255).

As shown in Chapter 7, centres that diagnosed and managed more patients were more likely to have access to all diagnostic techniques and provide treatment options that have been deemed effective for PCD patients by expert consensus statements. The constant throughput of patients and a higher number of patients under care in larger centres leads to increased knowledge of the disease by the multidisciplinary team responsible for patient care. In CF, best performing centres were those that had more follow-up appointments with patients, conducted spirometry and airway culture at every clinic visit and for all age groups, and used intravenous antibiotics for longer periods of time (254). Highest lung function values in children were seen across CF centres that performed closer microbiology surveillance and prescribed more oral corticosteroids (268).

The adoption of a high-level approach to service delivery for PCD, and rare diseases in general, facilitates coordination of resources and expertise between specialist centres. The development of networks such as BEAT-PCD and ERN-LUNG facilitates this approach by fomenting collaboration, research and knowledge exchange between specialists with different backgrounds and from different countries. This top-down method can ensure standardising of service delivery for PCD patients, which in turn will lead to optimisation of care. In Chapter 7, we found significant variability in routine clinical practice among PCD centres.

Structuring delivery of care for PCD patients around the expected number of cases of the disease will likely favour centralisation of service delivery; however, cultural and socio-economic factors should be considered. Social and cultural influences are considerable obstacles, particularly across EU Member States, as countries with relatively small populations like Belgium have clear divides along language and cultural heritage. Anecdotal accounts from PCD experts suggest that many French-speaking Belgium patients preferred to seek medical care in France than to travel to Flemish-speaking Leuven prior to the establishment of the PCD centre in Liege. This highlights the importance of contextual evidence when organising delivery of care for this rare disease.

Patient organisations can have a central role on establishing new PCD specialist centres as they can drive structural changes through engagement with policy makers. Most countries included in the survey detailed in Chapter 7 had an established PCD support group.

8.6 Strengths and limitations

I discussed the strength and limitations of the four studies included in this thesis in each of the corresponding chapters (please see 4.4.4, 5.5, 6.4 and 7.3.5 for more details). In this section, I will provide a summary of strengths and limitations that can be applied across the thesis.

8.6.1 Strengths

In this thesis I investigated delivery of care for PCD, from diagnostic methods to disease management and organisation of care. I used several different methods to address relevant questions in PCD research. One of the biggest strengths of this approach is that the techniques applied throughout this thesis are transferable to other rare diseases, as methodological challenges are similar across different rare diseases. In Chapter 4, I used disproportional sampling to increase the positive cases needed to conduct powered statistical analysis without requiring an unachievable number of cases. Chapter 5 shows the application of a novel data-driven approach that can be used to generate research hypotheses. By developing a discover and a validation topological model, I demonstrated that machine learning can also be applied to rare diseases.

The limited number of patients affected by the disease is a significant challenge in rare disease research in general. Collaboration across centres and with PCD centres in other countries is therefore inevitable in order to increase sample sizes and allow for powered and robust statistical analyses. In Chapter 4 we presented findings from the first powered study to investigate the accuracy of HSVA; previous studies did not perform sample size calculations (19, 98). We reported results from the largest genotype-phenotype study to date in Chapter 5, with almost 400 PCD patients included.

We were able to obtain large sample sizes due to large international networks. Members of the BEAT-PCD network were involved in all four projects presented in this thesis. We gathered data from three countries to explore genotype-phenotype associations, as described in Chapter 5, and we used the BEAT-PCD network to circulate the service delivery service, which was reported in Chapter 7. In fact, the unusually high response rate noted in Chapter 7 was a consequence of participation in the BEAT-PCD network.

Another strength of this thesis was the application of standardised data collection. In Chapter 4, all three participating centres used a data collection form and data manual for variable coding (see Appendix A). Similarly, I developed and circulated the data collection form and data manual for the variables used to generate the TDA models described in Chapter 5 in order to ensure standardised data collection. Results from the scoping review shown in Chapter 6 also benefited from a data collection proforma, developed a priori, to present results in a consistent manner. The use of this detailed form for data collection ensured that the high volume of data extracted by two different reviewers, each of whom extracted data from half of the studies deemed relevant for inclusion in the review, were consistent.

8.6.2 Limitations

The lack of a single test that can accurately and reliably diagnose all cases of PCD is a significant limitation to all studies. Selection bias might be an issue as cases that are easier to diagnose (*e.g.* those with mutation in known bi-allelic gene) have a higher chance of being included in studies since they are considered confirmed cases according to current evidence-based guidelines (25). This might be the case in both Chapter 4 and Chapter 5 and in many of the studies included in the scoping review presented in Chapter 6. However, there are many diseases that rely on combination testing or have multiple diagnostic criteria, none of which applied universally across all studies (*e.g.* asthma).

Two studies included in this thesis used routinely collected clinical data. The secondary use of data presents significant challenges, such as re-coding of data to fit study-specific definitions. This

is particularly poignant when pooling data from different PCD centres, as local differences in data collection and coding might have a considerable impact. As mentioned in 8.6.1, we used standardised data collection form for all studies included; however, centres collected and coded their own data, leaving room for differences in interpretation of previously coded data and the study data manual. For example, ethnicity coding varies considerably between countries, as detailed in 2.6. We could not find a universally accepted ethnicity code and used an adapted version of the Office of National Statistics codes, which are commonly used in research, particularly in the UK, but surprisingly not by NHS England. We therefore had to re-code ethnicity for all patients from UK sites, along with those from our collaborators, who collected ethnicity data as free text.

A common challenge in cross-sectional and retrospective studies such as the ones included in this thesis is the ability to minimise and handle missing data. In Chapter 5 we used a novel method that is robust to missingness in order to overcome this limitation. However, most PCD studies conducted to date do not provide details on how they handled missing data in their datasets.

The use of clinical and family history can lead to recall bias, particularly in patients that were diagnosed at an older age. Additionally, patients with more severe respiratory symptoms might not recall having milder upper airway symptoms, such as chronic rhinitis. This was explored in Chapter 5 as a possible explanation for the inverse association between upper and lower airway symptoms.

As discussed in Chapter 7, differences exist even between PCD centres from a single country. Table 2 in Chapter 4 depicts differences in equipment used to diagnose PCD across the three nationally commissioned PCD diagnostic services in England. It is unclear how these differences might impact sample processing and analysis.

As mentioned in 8.6.1, we benefited from large sample sizes. However, despite having more patients than most PCD research studies conducted to date, we were still limited in our analysis. For example, we could not perform multiple regression analysis in Chapter 7 to adjust for factors potentially contributing to differences seen across geographical regions and size of PCD centre. We were also unable to explore genotype-phenotype associations for other mutated genes aside from the two (*CCDC39* and *DNAH11*) selected for hypothesis testing, or investigate the impact of type of mutation, with or without complete or partial production of proteins, in Chapter 5.

8.7 Future directions for PCD clinical services

The implementation of new PCD centres and further development of existing centres should follow both the evidence-base from literature, such as diagnostic guidelines and management expert consensus, and the contextual evidence based on resources locally available, socio-economic and cultural background, and funding structures in place.

Policies established at higher levels, with the collaboration of groups of multinational researchers, clinicians and patient representatives, will likely lead to greater advances in delivery of care for patients with PCD. Research networks such as BEAT-PCD can guide clinical practice, particularly when most PCD experts participate in the development of guidelines and consensus statements, facilitating the acceptance and adoption of these into daily practice.

In terms of diagnostic testing, there has been a shift towards a more genetic-centric approach, since PCD is a genetic disease. However, it is important to highlight the additional contribution of other diagnostic tests such as HSVA and TEM, which are necessary in order to inform geneticists on the potential pathogenicity of observed mutations, particularly in new candidate genes and variants of unknown significance.

Disease management should be based on the adoption and measurement of disease-specific outcome measures that can be used to monitor patients throughout the years. These will allow for disease-specific management, as opposed to the current approach derived from CF and non-CF bronchiectasis management strategies.

8.8 Future research directions

There have been considerable recent advances in PCD knowledge; however, many aspects remain vastly unexplored. The research presented in this thesis provides starting points to various potential future projects.

There is still no universally accepted definition for CBP by HSVA, despite previous attempts to standardise nomenclature. Centres often use different categories to classify CBP. Recently, the TEM consensus statement provided criteria to classify ultrastructural defect by TEM into class 1 and class 2 defects and standardise reporting of findings (85). We need a similar consensus statement for CBP by HSVA, incorporating the views of scientists and researchers with vast experience in conducting HSVA, such as those involved in the study described in Chapter 4.

Current diagnostic guidelines rely on limited number of diagnostic accuracy studies. The lack of accuracy studies on genetic testing is in stark contrast to the current direction adopted by the PCD

community, with some centres favouring genetic testing as a first-line test over others. In Chapter 7, we showed that the proportion of centres conducting genetic testing in 2019 has greatly increased compared to those offering this technique in 2009. As new PCD genes are discovered at a rapid pace, studies examining the sensitivity, specificity, and positive and negative predictive values of genetic testing to diagnose PCD are needed.

There is a growing focus on the need for genotype-phenotype association studies (129). We used a novel method to explore the impact of different mutations on clinical and diagnostic phenotype in Chapter 5; however, we were unable to investigate the effect of type and location of mutations within a gene on phenotypic severity. Previous research has indicated that milder mutations that lead to the formation of proteins result in milder diagnostic phenotype (85, 178). This should also be explored in the context of disease severity, as it is possible that patients with these milder mutations also present with a milder disease severity.

This hypothesis was supported by our findings in Chapter 5, where *DNAH5*, the gene most commonly responsible for causing PCD, had a heterogenous presentation, with no distinct cluster in both the discover and validation models. The next step in genotype-phenotype association studies should be to test this hypothesis with well-designed and powered international studies. This will likely require large collaborations between European, North American and Australian PCD centres, as the number of patients with each type of mutation needed are considerably high. Focusing on a handful of common genes, such as *DHAN5, DNAH11, CCDC39, CCDC40, CCDC103,* and *DNAI1* might be a good initial strategy.

Importantly, the methods outlined in Chapter 5 could be used to generate important research hypothesis for future projects. For example, TDA can be used to investigate clinical decline in patients with different gene mutations in a longitudinal study. It could also be applied to explore correlations between pulmonary outcome measures and chronic colonisation with different respiratory pathogens.

An expert-led consensus with views of all stakeholders (e.g. clinicians, scientists, epidemiologists, patients and family members of patients with PCD) should be developed based on the data compiled and presented in Chapter 6 to recommend clinical outcome measures that should be collected routinely and in all future longitudinal cohorts and RCTs. We deliberately excluded ENT outcomes from our scoping review. These need to be investigated in a separate review, with or without a formal meta-analysis.

The study presented in Chapter 7 generated a significant amount of data that has not been fully analysed. Section C consists of questions on the opinions of PCD experts on various aspects of

service delivery for PCD. Additionally, other healthcare professionals and researchers from several countries were invited to complete section C of the survey to explore the opinions of different professionals involved in delivery of care. This, coupled with the rich qualitative data generated from the 15 interviews conducted to develop the survey, form the basis of a new mixed-methods project, with readily available data. As mentioned in Chapter 7, contextual data can potentially explain many of the variations seen across different service delivery models and this future project can explore some of those factors.

Furthermore, Chapter 7 underscored the need to investigate the impact of the different service delivery models on outcome measures, particularly those highlighted and recommended in Chapter 6. This will assist in the development of future clinical guidelines and new models for delivery of care for PCD, as variations in access to resources and contextual evidence can dictate the ability of each centre to adopt recommended diagnostic techniques, follow-up monitoring, and disease treatment. Additionally, it will allow for future benchmarking. A survival gap of approximately six years was observed between the seven CF care centres with the best health outcomes and the remaining centres across the Care Centre Network (269).

Studies examining different compositions of multidisciplinary team and optimal number of specialists in the PCD team are also needed. Data collected from the survey described in Chapter 7 provide a starting point for this type of research. Importantly, these studies should consider the contextual evidence, such as healthcare and funding structure, when determing the best strategy for delivery of care.

8.9 General conclusions

In this thesis, I addressed various aspects of delivery of care for patients with PCD, from accuracy and reliability of a widely used diagnostic test, correlations between genetic mutations and diagnostic and clinical phenotype, use of clinical outcome measures to monitor disease progression, and differences in service delivery across and within countries.

Chapter 4 showed that HSVA is accurate and reliable to diagnose PCD when done by experienced scientists. Findings from our study could strengthen the position of HSVA in the diagnostic pathway and can inform future versions of evidence-based diagnostic guidelines. We also highlight the importance of developing expert-led consensus statements to standardise the classification of beat patterns and the reporting of HSVA results.

Since PCD is a genetic disease, genotype-phenotype association studies are fundamental to better understand disease expression and progression. Chapter 5 showed that patients with *CCDC39*

mutations presented with a more severe respiratory phenotype at diagnosis, while those with mutations in *DNAH11* had less severe lower airway compromise. These findings could inform management of patients with these mutations, including frequency of follow-up appointments and aggressive treatment. The impact of different types of mutations (*e.g.* leading to production of protein or complete absence of protein) and location of mutations within the genes needs to be investigated.

The scoping review presented in Chapter 6 revealed a variety of clinical outcome measures that have been used in PCD research. However, these have been applied in an unstandardised manner, with studies using a variety of scales (*e.g.* CF adapted HRCT scoring systems) or measures (e.g. FEV_1 % predicted vs FEV_1 z-scores) to report their findings. This heterogeneity led to findings that are largely incomparable and to contradictory results. Chapter 6 underscored the need to develop a core set of clinical outcome measures for disease monitoring in future studies, particularly prospective cohorts and RCTs. Spirometry, HRCT scans, HRQoL and the application of standardised disease-specific definition of pulmonary exacerbations should be used across PCD research studies in order to generate comparable results.

Finally, Chapter 7 compared differences in service delivery models across geographical regions and size of PCD centre. Comparisons with a survey conducted ten years ago revealed that service delivery has become more centralised and standardised across European countries, particularly diagnostic testing after the advent of the evidence-based PCD diagnostic guidelines published by the ERS in 2017. In terms of management strategies, larger centres tended to adopt measures centred around airway clearance (*e.g.* respiratory physiotherapy, formal exercise programmes and nasal rinsing) and immunisation, while smaller PCD centres were still favouring therapies that are commonly used in other diseases but have no apparent benefit in PCD (*e.g.* bronchodilators and inhaled corticosteroids).

The findings from the studies included in this thesis will inform future research and further our understanding of diagnosis and clinical management of PCD patients.

Appendix A Accuracy of high-speed video analysis to diagnose PCD

A.1 Variable coding

Clinical and diagnostic variables were extracted and coded using a standardised data collection form, agreed a priori (Table 15).

Table 15. Data dictionary for clinical and diagnostic data included in the study depicted in Chapter 4.

Variable	Categories
Study ID	*numerical/assigned by Study coordinator
Month of birth	*numerical
Year of birth	*numerical
Gestational age	1: Pre-term (<37 weeks)
	2: Full term (>= 37 weeks)
Chest symptoms in neonatal period	0: No
	1: Yes
	2: Unknown
Admission to neonatal unit	0: No
	1: Yes
	2: Unknown
Situs abnormality (situs inversus or heterotaxy)	0: No
	1: Yes
	2: Unknown
Congenital heart defect	0: No

Variable	Categories
	1: Yes
	2: Unknown
History of persistent rhinitis	0: No
	1: Yes
	2: Unknown
History of chronic ear or hearing symptoms (e.g. glue ear, serious otitis media, hearing loss, ear perforation)	0: No
	1: Yes
	2: Unknown
nNO measurement date	*date
nNO results (in ppb)	*numerical Highest mean, if done in more than one occasion
nNO analyser	Free text
nNO flow	*numerical
nNO measurement method	1: Velum closure
	2: Tidal breathing
	3: Other (specify as free text)
Ambient nitric oxide	*numerical
TEM date	*date
TEM result	1: Normal
	2: Outer dynein arm
	3: Inner dynein arm + microtubular disarrangement
	4: Inner dynein arm

Variable	Categories
	5: Inner and outer dynein arm
	6: Microtubular disarrangement
	7: Central complex defect
	8: Lack of cilia
	9: Inconclusive
	0: Not done
	Comments (free text)
Genetic test date	*date
Gene name	Free text
Pathogenicity score (if available)	1, 2, 3, 4, 5
Multidisciplinary team decision	1: PCD positive
	2: PCD highly likely
	3: PCD highly unlikely
	4: Inconclusive
Multidisciplinary team inconclusive	If multidisciplinary team decision was inconclusive, detail reason given (free text)
Multidisciplinary team decision comment	Comment on diagnostic tests used to reach a final decision - list all tests considered during multidisciplinary team meeting e.g. nNO, HSVA, TEM, immunofluorescence results, computed electron tomography, cell culture (Free text)
ERS classification (multidisciplinary team decision reclassified	PCD Positive
	PCD highly likely
	PCD highly unlikely

Variable	Categories
according to ERS guidelines)	Inconclusive

*nNO: Nasal nitric oxide, TEM: transmission electron microscopy, HSVA: high-speed video analysis, ERS: European Respiratory Society, PCD: primary ciliary dyskinesia

A.2 Variable definition

Table 16. Data dictionary for outcome variables included in the study depicted in Chapter 4

Outcome Variables	Categories	Definitions
Final outcome for each patient-sample	PCD positive	PCD positive with clear abnormalities seen on HSVA. Further investigations may be needed to support and to phenotype/genotype disease
	PCD highly likely	Abnormal HSVA outcomes but not as clear to classify it as 'PCD positive'. Further investigations may be needed to confirm PCD and to phenotype/genotype disease
	PCD highly unlikely	PCD highly unlikely due to a normal HVSA outcome. Samples for further investigation to be stored (clinical history or nNO might override this decision)
	Inconclusive	Due to severe secondary infection/ inflammation, a lack of cilia, or an inadequate sample yield. Repeat investigations are required and cell culture is desirable along with other tests if possible
Ciliary beat pattern overall decision	Normal	Majority of cilia demonstrate a complete power and recovery stroke on at least six strips of healthy ciliated epithelium
	Completely immotile	Absent ciliary movement on all strips viewed (except in response to flow, microscope shaking, etc) usually seen in inner dynein arm and outer dynein arm defects

Outcome Variables	Categories	Definitions
	Weak residual movement	Majority of cilia have a complete absence of movement; however, occasional cilia perform a slow or stiff beat. As usually seen with *DNAH5* defects
	Stiff	The cilia lack bending action. Incomplete power or recovery stroke on most cilia viewed
	Rotating	Majority of cilia circling, easiest observed from the top view. As those with a central complex/transposition defect
	Staggered beat	Some cilia seen pausing, interrupted beat or double beating, as seen in those with *CCDC164* and *CCDC65* mutations
	Long with bulbous tips	Some cilia in the sample appear long (approximately twice length of the other cilia), occasionally with a whip-like flagella movement, with occasional accumulations or swollen ciliary tips. Sometimes resulting in the long cilia becoming static or bent over
	Lack of cilia	No cilia throughout the sample or <10 cilia per cell on otherwise apparently healthy differentiated ciliated epithelium
Ciliary beat pattern side view predominant finding	Same as above	

Outcome Variables	Categories	Definitions
CBP top view predominant finding	Same as above	
If stiff report location	Apical	Base of the cilia moves well but the top does not have a full range of movement (sometimes forming a 'question mark' type shape)
	Basal	Limited movement in the region of the cilia concordant with the microvilli - as in a *DNAH11* defect
	Global	Limited bend throughout the length of the cilia. Scissor like motion as in *CCDC39* and *CCDC40* (microtubular disorganisation defects)
	Not applicable	
Synchronisation of ciliary beat pattern present	Yes	Most cilia are beating in the same direction. You can picture which direction mucus would be transported
	No	The opposite of the above. Cilia are beating in different directions
	Mixed	There are ≥2 predominant ciliary beat patterns, which may include both synchronous and asynchronous beating

Outcome Variables	Categories	Definitions
	Not applicable	
Mucociliary clearance visualised	Yes	Blood or debris within the sample were seen on at least one occasion being swept by the cilia in the direction of ciliary beat
	No	No particles within the sample were seen on at least one occasion being swept by the cilia in the direction of ciliary beat
	Not applicable	Referring to samples with static or no cilia, or insufficient samples
Mucus contamination present	0	No mucus
	1	Minor quantity of mucus visualised, no ciliary impedance
	2	Moderate quantity of mucus visualised, some ciliary impedance (can be avoided)
	3	Severe mucus, unavoidable ciliary impedance
Bacterial contamination present	Yes	Bacteria visualised in at least one clip
	No	No or few bacteria visualised
	Yes	Red blood cells visualised in at least one clip

Outcome Variables	Categories	Definitions
Blood contamination	No	No blood seen
Cell health compromised	Yes	On one strip or more the cell membrane is extruding or discontinuous, or only small clusters of cells are seen, or several inflammatory cells are seen
	No	All strips have more than 10 cells and a continuous apical membrane

A.3 Additional results

Table 17. Diagnostic outcomes decisions made by each of the three high-speed video analysis (HSVA) scientists (*i.e.* scientist A-C, *n*=120 for each) compared to the ERS guidelines (25).

	PCD positive* by ERS	**PCD highly unlikely by ERS**	**PCD highly likely by ERS**	**Inconclusive by ERS**	**Total**
Outcomes by scientist A					
PCD positive	32	0	7	4	43
PCD highly unlikely	0	21	2	13	36
PCD highly likely	2	0	4	2	8
Inconclusive	2	5	3	23	33
Total	36	26	16	42	120
Outcomes by scientist B					
PCD positive	35	2	12	8	57
PCD highly unlikely	0	8	0	8	16
PCD highly likely	1	4	2	9	16
Inconclusive	0	12	2	17	31
Total	36	26	16	42	120

	PCD positive* by ERS	PCD highly unlikely by ERS	PCD highly likely by ERS	Inconclusive by ERS	Total
Outcomes by scientist C					
PCD positive	27	0	6	1	34
PCD highly unlikely	0	24	2	21	47
PCD highly likely	7	0	5	6	18
Inconclusive	2	2	3	14	21
Total	36	26	16	42	120

ERS: European Respiratory Society; HSVA: high-speed video analysis; PCD: primary ciliary dyskinesia; *'PCD positive' cases were those with a hallmark TEM defect and/or confirmed mutation on PCD gene.

Table 18. Diagnostic outcomes decisions made by each of the three high-speed video analysis (HSVA) scientists (*i.e.* scientist A-C, *n*=120 for each) compared to the original diagnostic decision made by the MDT

	PCD positive* by expert MDT	PCD highly unlikely by expert MDT	Inconclusive by expert MDT	Total
Outcomes by scientist A				
PCD positive*	49	0	2	51
PCD highly unlikely	2	26	8	36
Inconclusive	8	10	15	33
Total	59	36	25	120

Appendix A

	PCD positive* by expert MDT	**PCD highly unlikely by expert MDT**	**Inconclusive by expert MDT**	**Total**
Outcomes by scientist B		y		
PCD positive*	55	7	11	73
PCD highly unlikely	0	14	2	16
Inconclusive	4	15	12	31
Total	59	36	25	120
Outcomes by scientist C				
PCD positive*	47	0	5	52
PCD highly unlikely	2	33	12	47
Inconclusive	10	3	8	21
Total	59	36	25	120

PCD: primary ciliary dyskinesia; MDT: multidisciplinary team; HSVA: high-speed video analysis; * Includes both 'PCD positive' and 'PCD highly likely' outcomes

Table 19. Diagnostic outcomes decisions from a random subset of patient-samples (*n*=20) by each of the three scientists (*i.e.* scientist A-C) using high-speed video analysis (HSVA) at two different time points (*i.e.* original study analysis and re-analysis after one year)

	PCD positive re-assessment by scientist A	**PCD highly unlikely re-assessment by scientist A**	**Inconclusive re-assessment by scientist A**	**Total**
Original outcome by scientist A				
PCD positive	7	0	0	7
PCD highly unlikely	0	4	3	7
Inconclusive	1	0	5	6
Total	8	4	8	20
	PCD positive re-assessment by scientist B	**PCD highly unlikely re-assessment by scientist B**	**Inconclusive re-assessment by scientist B**	**Total**
Original outcome by scientist B				
PCD positive	10	0	0	10
PCD highly unlikely	0	2	1	3
Inconclusive	2	1	4	7
Total	12	3	5	20

Appendix A

	PCD positive re-assessment by scientist C	**PCD highly unlikely re-assessment by scientist C**	**Inconclusive re-assessment by scientist C**	**Total**
Original outcome by scientist C				
PCD positive	6	0	0	6
PCD highly unlikely	0	5	2	7
Inconclusive	1	0	6	7
Total	7	5	8	20

PCD: primary ciliary dyskinesia

Appendix B Previously described genotype-phenotype associations in PCD

Table 20. Summary of the phenotype-genotype relationships described in the PCD literature.

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Dynein structure	*DNAH5*	Dynein heavy chain	ODA subunit	ODA (class 1)	Predominantly static	Randomised L-R asymmetry (37)
Dynein structure	*DNAH9*	Dynein heavy chain	ODA subunit	ODA (class 1)	Normal/stiff at top	Randomised L-R asymmetry Mild or no respiratory involvement (272)
Dynein structure	*DNAH11*	Dynein heavy chain	ODA subunit	Normal (class 2)	Hyperfrequent and stiff at base, with reduced amplitude	Randomised L-R asymmetry (36)
Dynein structure	*DNAI1*	Dynein intermediate chain	ODA subunit	ODA (class 1)	Predominantly static	Randomised L-R asymmetry
Dynein structure	*DNAI2*	Dynein intermediate chain	ODA subunit	ODA (class 1)	Predominantly static	Randomised L-R asymmetry
Dynein structure	*DNAL1*	Dynein light chain	ODA subunit	ODA (class 1)	Predominantly static	Randomised L-R asymmetry

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Dynein structure	*NME8 (TXNDC3)*	Dynein light chain	ODA subunit	ODA/variable (normal) (class 1)	Predominantly static	Randomised L-R asymmetry
Dynein assembly	*CCDC103*	Coiled-coil domain	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Variable, ranging from normal to areas of immotile cilia	Lower rate of bronchiectasis and increased chance of OM. Higher levels of nNO (35) Randomised L-R asymmetry
Dynein structure	*CCDC114*	Coiled-coil domain	ODA targeting/docking centre	ODA (class 1)	Abnormal, large areas of immotile cilia with some residual movement	Randomised L-R asymmetry , Fertile (8, 53)
Dynein structure	*CCDC151*	Coiled-coil domain	ODA targeting/docking centre	ODA (class 1)	Predominantly static	Randomised L-R asymmetry
Dynein structure	*ARMC4*	Armadillo repeat	ODA targeting/docking centre	ODA (class 1)	Predominantly static	Randomised L-R asymmetry
Dynein structure	*TTC25*			ODA(class 1)	Predominantly static	Randomised L-R asymmetry

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Dynein assembly	*DNAAF1 (LRRC50)*	Cytoplasmatic	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry Infertility reported (53)
Dynein assembly	*DNAAF2 (KTU)*	Cytoplasmatic	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Some residual movement	Randomised L-R asymmetry Motile defects in the sperm
Dynein assembly	*DNAAF3*	Dynein assembly factor	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry Infertility reported (273)
Dynein assembly	*DYX1C1 (DNAAF4)*	Dynein assembly factor	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static or some residual movement	Randomised L-R asymmetry Infertility reported (274)
Dynein assembly	*DNAAF5 (HEATR2)*	Dynein assembly factor	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Dynein assembly	*SPAG1*	Sperm-associated antigen 1	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry
Dynein assembly	*LRRC6*	Protein TILB homolog	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry Fertile (53)
Dynein assembly	*C21ORF59*		Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry
Dynein assembly	*ZMYND10*	Zinc finger MYND domain	Cytoplasmatic dynein arm assembly or transport factor	ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry
Dynein assembly	*CFAP300*			ODA+IDA	Completely static	Randomised L-R asymmetry Motility defects in the sperm (275)
Dynein assembly	*PIH1D3*	X-linked		ODA+IDA (class 1)	Completely static	Randomised L-R asymmetry

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Radial spoke or central complex	*HYDIN*	Hydrocephalus-inducing protein homolog	Central pair subunit	Normal/CC defect (class 2)	Rotation/stiffness	Situs solitus (118)
Radial spoke or central complex	*STK36*		Central pair subunit	Normal/CC defect (class 2)	Dyskinetic	Situs solitus
Radial spoke or central complex	*RSPH1*	Radial spoke head	Radial spoke head subunit	Intermittent CC defect (class 2)	Rotation/stiffness	Situs solitus, milder lung disease, lower rates of NRDS, increased chance of OM, infertility. Higher levels of nNO (48, 78)
Radial spoke or central complex	*RSPH3*	Radial spoke head	Radial spoke head subunit	Intermittent CC/MTD defect (class 2)	Rotation/stiffness	Situs solitus, increased chance of OM (48)
Radial spoke or central complex	*RSPH4A*	Radial spoke head	Radial spoke head subunit	Intermittent CC defect (class 2)	Rotation/stiffness	Situs solitus, increased chance of OM, fertile (48, 53)

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Radial spoke or central complex	*RSPH9*	Radial spoke head	Radial spoke head subunit	Intermittent CC defect (class 2)	Rotation/stiffness	Situs solitus, increased chance of OM (48)
Radial spoke or central complex	*DNAJB13*			Intermittent CC defect (class 2)	Rotation/stiffness	Situs solitus
N-DRC/ molecular ruler	*CCDC164 (DRC1)*	Dynein regulatory complex	Nexin link subunit	Normal/MTD (class 2)	Normal/stiffness	Situs solitus
N-DRC/ molecular ruler	*CCDC65 (DRC2)*	Dynein regulatory complex	Nexin link subunit	Normal/MTD (class 2)	Normal/stiffness	Situs solitus
N-DRC/ molecular ruler	*GAS8 (DRC4)*	Dynein regulatory complex	Nexin link subunit	Normal/MTD (class 2)	Normal/stiffness	Situs solitus

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
N-DRC/ molecular ruler	*CCDC39*	Dynein regulatory complex	Axonemal ruler	IDA+MTD (class 1)	Variable, ranging from stiffness to immotility	More severe lung disease, higher rate of bronchiectasis, worse weight gain and infertility (28) Randomised L-R asymmetry
N-DRC/ molecular ruler	*CCDC40*	Dynein regulatory complex	Axonemal ruler	IDA+MTD (class 1)	Variable, ranging from stiffness to immotility	More severe lung disease, higher rate of bronchiectasis, worse weight gain and infertility (28) Randomised L-R asymmetry
Other functions	*LRRC56*	Intra-flagella transport-dependent delivery of ODA components		Subtle ODA/Normal (class 2)	Dyskinetic	Randomised L-R asymmetry
Other functions	*RPGR*	X-linked	Function related to non-motile-cilia	Variable		Situs solitus
Other functions	*OFD1*	X-linked	Function related to non-motile-cilia	Unknown		Randomised L-R asymmetry

Gene group	Gene	Protein (86)	Axonemal/cellular structure or function (270, 271)	TEM (class defect) (85)	Ciliary beat pattern	Clinical phenotype
Other functions	*CCNO*	Cyclin-O	Cytoplasmatic centriole assembly and docking factor	Severely reduced cilia, mislocalisation of basal bodies in the cytoplasm (class 2)	Reduced cilia	Situs solitus, more severe lung disease and bronchiectasis and at younger age (177) Possible fertility problems (276, 277)
Other functions	*MCIDAS*	Multicilin	Nuclear regulator of *CCNO* and *FOXJ1*	Severely reduced cilia (class 2)	Reduced cilia	Situs solitus, more severe lung disease and bronchiectasis and at younger age, hydrocephalus reported (176)

TEM: transmission electron microscopy; ODA: outer dynein arm; L-R: left-right; IDA: inner dynein arm; OM: otitis media; nNO: nasal nitric oxide; CC: central complex; MTD: microtubular disorder; NRDS: neonatal respiratory distress syndrome.

Appendix C Clinical outcome measures for use in PCD: studies included in the scoping review

C.1 Study characteristics

Table 21. Characteristics of studies included in the clinical outcome measures scoping review

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Abitbul *et al* (2016) (259)	Israel	2012 to 2013	Retrospective cohort (?)	150	N/A	Mean (SD): 17.08 (11.96), median: 15.05, range: 0.15 to 60.47	At least one of the following: nNO + HSVA, TEM, IF or genetic testing	Inclusion: clinical symptoms consistent with PCD phenotype. Exclusion: acute respiratory infection 4 weeks prior to study
Ahmad *et al* (2015) (237)	UK	January 2008 to May 2014	Retrospective study	19	Healthy controls (17)	Median: 13.89	Not reported	Not reported
Alanin *et al* (2015) (278)	Denmark	January 2002 to December 2012	Retrospective cohort	107	N/A	Median: 17, range 0 to 74	Clinical symptoms +	Definitive PCD diagnosis + microbiology data available

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
							(TEM, HSVA or genetic testing)	
Behan *et al* (2017) (225)	UK, USA, Canada	Between April 2014 and March 2016	Mixed cross-sectional and longitudinal study (for 10 participants that were re-assessed during an exacerbation)	72	N/A	Mean (SD): 34.8 (17.3) for UK, range 18 to 79; 31 (12.9) for USA/ Canada, range: 18 to 65	UK participants: clinical phenotype + HSVA and/or TEM. North American participants: clinical phenotype + TEM and/or genetic testing.	Adults (aged ≥18 years) with diagnosis of PCD in one of the specified diagnostic centres and ability to read and speak English fluently.
Boon *et al* (2014) (279)	Belgium	Jan 1990 to August 2012	Retrospective study	168	N/A	Median (IQR): 17.7 (9.5 to 28.1)	(HSVA or TEM) + cell culture	Not reported

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Boon *et al* (2015) (280)	Belgium	May 2011 and September 2014	Prospective observational study	38	Healthy controls (70)	Median (IQR): 16.1 (11.1 to 19.6)	HSVA + cell culture	Inclusion: chest HRCT within 1 year of the MBW measurement, and without exacerbations Exclusion: history of prematurity, asthma, allergy or recurrent respiratory symptoms
Bush *et al* (2006) (281)	UK	Not reported	Unclear	19	CF children (30)	Mean (SD): 9.5 (3)	nNO, CBF and TEM	Not reported
Carotenuto *et al* (2013) (282)	Italy	December 2011 to September 2012	Cross-sectional questionnaires	10	Healthy children and adolescents (34)	Range: 6 to 16	nNO, HSVA and TEM	Exclusion: upper and lower respiratory tract infection and asthma exacerbation, heart disease, mental retardation (IQ less than

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
								70), epilepsy, and psychiatric disorders
Cockx *et al* (2017) (283)	Belgium	2012 to 2016	Case-control	36	Healthy controls (40); 21 children and 19 adults	Mean: 13, range 2 to 26	HSVA, cell culture, TEM, genetic testing	Clinically stable
Cohen-Cymberknoh *et al* (2017) (284)	Israel, Belgium, Italy, Germany	January 2008 to December 2013	Retrospective study	217	N/A	Median (SD) 19.9 (13.9), range 0 to 67	According to European consensus (265)	Patients with follow-up data for at least 3 years + results from at least 2 sputum cultures
Cohen-Cymberknoh *et al* (2014) (105)	Israel	2007 to 2011	Cross-sectional study	34	CF patients (130); CF-PI (88), CF-PS (42)	Mean (SD): 15.9 (8.6)	Clinical phenotype + ((nNO + TEM), HSVA, genetic testing)	Confirmed diagnosis of PCD or CF + available spirometry, HRCT, sputum cultures and pancreatic sufficiency test

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Davis *et al* (2015) (285)	USA, Canada	2006 to 2012	Cross-sectional study	118	N/A	Median (unclear): 8, range 5 to 11	TEM or genetic testing	<19 years of age and confirmed diagnosis of PCD
Ellerman *et al* (1997) (124)	Denmark	Late 1970s to 1994, with minimum of 2 years follow-up	Prospective cohort	24	N/A	Median 21, range 2 to 56	Clinical phenotype + HSVA + normal sweat test to exclude CF	Inclusion: confirmed diagnosis + regular spirometry. Exclusion: CF patients
Frija-Masson *et al* (2017) (122)	France	1990 to 2010	Retrospective cohort	78	N/A	Median (IQR): 34.8 (28.6 to 47.1), range 18 to 77	Clinical phenotype or TEM or genetic testing	Not reported
Gokdemir *et al* (2014) (232)	Turkey	Not reported	Randomised controlled	24	N/A	Mean (SD): 12.9 (2.7),	Clinical phenotype or TEM	Inclusion: clinical stability Exclusion: history of pneumothorax, massive

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
			crossover study			range 7 to 18		haemoptysis or congestive heart failure
Green *et al* (2012) (112)	Denmark	Not reported	Cross-sectional prospective study (?)	27	N/A	Median: 11.3, range 6.3 to 18.5	Clinical phenotype + nNO, HSVA, TEM. CF and immunodeficiency were excluded	Patients <=18 years diagnosed with PCD + stable clinical condition on day of MBW measurement
Green *et al* (2016) (231)	Denmark	Not reported	Cross-sectional prospective study (?)	28	CF (61) and healthy controls (48)	Median (IQR): 12.4 (10.7 to 14.6)	According to consensus guidelines (265)	Diagnosed CF or PCD, age from 5 to 18 years; healthy controls without chronic or recurrent lung disease, fever, or symptoms of respiratory tract infection in the previous 4 weeks

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Hellinckx *et al* (1998) (286)	Belgium	1996	Longitudinal study, no further details provided (?)	12	N/A	Mean (SD): 15.2 (7.0), range 6 to 32	Clinical phenotype + HSVA and TEM	Patients with PCD in regular follow-up for 3 to 20 years
Irving *et al* (2013) (118)	UK	Not reported	Case-control (?)	33	CF patients (127)	Mean: 24.66 Mean for subgroup of 21 PCD for HRCT: 31.2	According to Bush *et al* (287).	Not reported
Irving *et al* (2017) (230)	UK	2009 to 2010; 2014 to 2015	Prospective cohort	29	N/A	Median: 14, range 3 to 53	TEM or genetic testing	Not reported
Jain *et al* (2007) (216)	UK	Not reported	Retrospective study	89	N/A	Median: 4, range 0 to 14.4	nNO, LM, TEM + tests to exclude CF and immunodeficiency	Not reported

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Joensen *et al* (2014) (209)	Denmark	May 2013 to September 2013	Cross-sectional case-control study (?)	21	CF patients (64) and healthy controls (21)	Median (IQR): 26.0 (19.0 to 45.5)	Clinical symptoms + abnormal ciliary beat pattern + TEM	Exclusion for controls: active use of tobacco or a history of pulmonary disease, inflammatory disease, metabolic, or genetic disorders; fever or productive coughing 14 days prior to measurement
Kawakami *et al* (1996) (288)	Japan	Not reported	Cross-sectional questionnaires	48	N/A	Mean (SE): 38.4 (1.7), range 17 to 72	Clinical symptoms and/or TEM	Not reported
Kennedy *et al* (2007a) (289)	USA	August 2003 to March 2006 for prospective	Mixture of prospective and	142	N/A	Mean (SD) for *n*=7 with outcome	TEM (only reported for *n*=7)	Not reported

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
		study; Prior to August 2003 for retrospective study	retrospective study			measure: 56 (7)		
Kennedy *et al* (2007b) (119)	USA	January 1995 to May 2006	Retrospective cross-sectional (?)	45	N/A	Mean (SD): 29 (3)	Clinical phenotype + TEM, nNO	Chest CT available from cohort of 140 PCD patients (289)
Knowles *et al* (2014) (285)	USA	Not reported	Cross-sectional study	90	N/A	Mean (SD): 35.3 (18.6) (RSPH1 mutations) Mean (SD): 34.2 (17.6) (75 age- and	TEM or genetic testing	Not reported

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
						sex matched)		
Koh *et al* (2000) (108)	South Korea	Not reported	Randomised double-blinded, placebo-controlled, cross-over study	19	N/A	Median: 12, range 7 to 16	TEM	Children that could perform spirometry
Li *et al* (2005) (290)	UK	1986 to 2002	Retrospective study	20	N/A	Not reported	Clinical phenoytpe, nNO (14% of cases), LM (49% of cases), TEM (70% of cases)	Inclusion: HRCT-diagnosed bronchiectasis in subjects with suggestive clinical features Exclusion: CF diagnosed by sweat test and/or

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								analysis of genetic testing.
Loomba *et al* (2017) (291)	USA (isomerism patients and healthy control) and Denmark (PCD patients)	January 1998 to December 2014	Retrospective case-control study (?)	17	Healthy controls (17), patients with Fontan + isomerism (17), patients with Fontan - isomerism (17)	Mean (SD): 13.36 (3.5)	Not reported, but used the same cohort as Madsen *et al* (292)	Not reported
Lopes *et al* (2015) (215)	Brazil	Not reported	Cross-sectional study	11	Tuberculosis patients (34), non-tuberculosis infection (29), CF	Mean (SD): 56 (18.7)	Clinical phenotype + TEM	Inclusion: individuals with bronchiectasis based on HRCT findings, clinically stable, no

Appendix C

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
					(21), rheumatoid arthritis (17)			history of smoking, and >=18 years of age. Exclusion: history or diagnosis of asthma (n= 18) or a pleural (n= 10) or cardiovascular disease; subjected to lung resection (n= 4) or used oral corticosteroids 4 weeks before the study; unknown cause of bronchiectasis (n= 35); traction bronchiectasis secondary to interstitial lung disease.

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Madsen *et al* (2013) (292)	Denmark	Not reported	Case-control study	44	Healthy controls (33)	Median (IQR): 14.8 (6.5 to 29.7)	Clinical phenotype, nNO (n=42), HSVA (n=42), TEM (n=39) (265)	Inclusion: children and young adults; healthy age-, gender- and BMI-matched non-atopic subjects with normal spirometry as controls. Exclusion: unable to perform pulmonary function testing or exercises (e.g mental or physical disability or known cardiovascular disease)
Maglione *et al* (2012) (293)	Italy	2007 to 2010 (?)	Retrospective cohort study	20	N/A	Median: 11.6, range 6.5 to 27.5	HSVA, TEM	Inclusion: availability of CT scan and spirometry at some time point during the follow-up in

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								stable patient, and of a second CT scan plus spirometry during exacerbation. Exclusion: < 6 years of age, unable to perform spirometry, or had only one CT scan during follow-up
Maglione *et al* (2014a) (294)	UK, Italy and Denmark	UK: 1990 to 2011 Denmark: 1979 to 2011 Italy: 1994 to 2011	Cross-sectional and longitudinal study (?)	158	N/A	Median at first spirometry: 8.7, range 4.2 to 17.4	TEM	Ability to perform reliable spirometry, and availability of annual anthropometric and spirometry data over the last 3 years

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Maglione *et al* (2014b) (208)	Italy	Not reported	Prospective questionnaire	20	N/A	Median: 16.9; range 12 to 33.4	Not reported	Not reported
Maglione *et al* (2017) (295)	Italy	January 2014 to May 2015	Prospective, single-center	20	CF patients (20)	Median: 15.1, range 8.7 to 29.4	nNO, HSVA, TEM, genetic testing	Mild CF patients: selected according to the functional criteria described by Schluchter *et al* (296). PCD patients: stable lung disease, without acute dyspnea or cough, no pulmonary function changes and no requirement for intravenous antibiotics in the previous 4 weeks. Exclusion: acute

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
								respiratory infection, developmental delay, or other conditions that could compromise compliance to MRI or spirometry e.g. age < 6 years, claustrophobia.
Magnin *et al* (2012) (221)	France	1988 to 2010	Retrospective cohort study	20	N/A	Median (IQR) at first visit: 4.7 (1.7 to 7.9), range 0 to 13.8	Clinical phenotype, HSVA, TEM, computerised EM (for IDA defects, after 2002)	Inclusion: age < 15 years at the beginning of follow-up, at least 8 years of follow-up, at least 2 concurrent CT and lung function tests available in a phase of clinical stability of the lung disease

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								without modification of the treatment regimen in the last 4 weeks.
Marthin *et al* (2010) (120)	Denmark	Late 1970s onwards	Partly cross-sectional and partly designed as an uncontrolled, observational, single-group, single-centre, longitudinal and retrospective study of prospectively collected data	74	N/A	Median at first visit (1979): 9, range 4.4 to 43.7	(Clinical phenotype + HSVA), (nNO, TEM, pulmonary radioaerosol mucociliary clearance) in most patients	Inclusion: at least 1.5 years of follow-up and acceptable spirometry Exclusion: uncertain diagnosis, unable to perform reliable spirometry and nonvalid LF measurements

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
McManus *et al* (2003) (297)	UK	January 2003 to April 2013	Cross-sectional (questionnaires)	93	N/A	Median 16.5 (IQR 10.8 to 31.3)	Not reported	Patients on the mailing list of the UK's PCD Family Support Group
McManus *et al* (2006) (298)	UK	Januray 2003	Cross sectional (questionnaires)	71	N/A	Median (IQR): 20.1 (15.6 to 38.7)	Not reported	Patients on the mailing list of the UK's PCD Family Support Group
Mirra *et al* (2015) (299)	Italy	March to June 2012	Prospective, cross-sectional study	22	N/A	Median: 10.5, range 2 to 34	HSVA, TEM	Inclusion: stable patients with confirmed diagnosis of PCD, according to Maglione *et al* (293) Exclusion: airway infections or asthma exacerbations during the 4 weeks prior to enrolment, current

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								smoker, long term use of oral steroids, antibiotic treatment in the last 4 weeks before enrolment, prescription of over-the-counter calcium or vitamin-D supplements prior to, or during the study period.
Montella *et al* (2009a) (229)	Italy	Not reported	Prospective, cross-sectional study	13	N/A	Median: 15.2; range 10.4 to 29.3	LM, TEM	Not reported
Montella *et al* (2009b) (300)	Italy	March 2007 to June 2008	Prospective, cross-sectional study	14	Primary immunodeficiency patients (14), recurrent pneumonia (13)	Median: 15.2, range 10.4 to 29.3	Clinical phenotype, LM, TEM	Inclusion: patients with PCD, chronic lung disorders, primary immunodeficiency, recurrent pneumonia

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
								Exclusion: acute respiratory infection and/or mental retardation or other conditions that could compromise compliance to HRCT and MRI (e.g. age <5 years, claustrophobia)
Montuschi *et al* (2014) (301)	Italy	Not reported	Cross-sectional study	45	Primary analysis: CF (21), age-matched healthy controls (21) Validation subjects: CF (25), age-matched	Mean (SD) primary Analysis: 17.4 (0.9), range 11 to 32 Mean (SD) validation	PCD and CF were diagnosed according to published criteria (302, 303)	Not reported

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
					healthy controls (25)	subjects: 15.7 (0.6), range 11 to 31		
Noone *et al* (1999) (233)	USA	Not reported	Double blind, randomised, crossover study	12	N/A	Mean: 34, range 14 to 71	TEM	Exclusion: significant intercurrent infection, defined as a change in cough or sputum production or increased dyspnea within 2 weeks of screening
Noone *et al* (2004) (38)	USA	1994 to 2002	Prospective cohort (unclear)	78	N/A	Mean: 26.8; median: 29, range 0 to 73	Clinical phenotype, nNO, HVSA, TEM	Exclusion: atypical asthma, CF, allergic bronchopulmonary aspergillosis, Young's Syndrome, and idiopathic bronchiectasis

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Nyilas *et al* (2017) (304)	Germany and Switzerland	March 2013 to April 2015	Cross-sectional multicentre study	49	37	Mean (SD): 14.7 (6.6), range 11 to 18	Clinical phenotype, HSVA + (TEM, IF or genetic testing)	Not reported
Oktem *et al* (2013) (305)	Turkey	Not reported	Cross-sectional study	29	29	Mean (SD): 10.0 (5.9), range 0.5 to 24	Clinical phenotype, TEM	Not reported
Olveira *et al* (2017) (210)	Spain	2002 to 2011	Multicenter, nested cross-sectional study from Spanish registry	60	Other causes of bronchiectasis (n = 1987)	Mean (SD): 42.9 (18.8)	Clinical phenotype, nNO, TEM, saccharin test and labelled seroalbumin for differential diagnosis	Adult patients with bronchiectasis

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Paff *et al* (2013) (306)	The Netherlands	August to November 2011	Cross-sectional case–control study	25	CF (25), healthy controls (23)	Median (IQR): 10.7 (7.1 to 14.5)	Clinical phenotype, HSVA, TEM (265)	Exclusion: children with any pulmonary, inflammatory or metabolic disease.
Paff *et al* (2017) (106)	The Netherlands	April 2014 to May 2015	Double blind randomised controlled crossover trial over a 28-week period with 4 weeks washout	22	N/A	Median (IQR): 47.6 (26.9 to 58.1)	Not reported	Inclusion: ≥ 18 years, clinically stable, FEV_1 had to be at least 40% of the predicted value for height, age and sex and within 10% of the best value obtained during the previous six months. Exclusion: women with a current or intended pregnancy or who were breastfeeding, cigarette smokers, known quinine

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								sulphate allergy, or in use of the following medications: hypertonic saline, rhDNAse, N-acetylcysteine or non-routine antibiotics in the previous 4 weeks. Participants whose oxygen saturation fell under 90% or whose FEV_1 fell more than 15% compared to its prebronchodilator value 15 minutes after inhalation of a test solution with hypertonic saline and taste-masking

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								agent, were not eligible to proceed in the trial.
Phillips *et al* (1998) (213)	UK	Not reported	Cross-sectional	12	12	Median: 11, range 7 to 15	Clinical phenotype, HSVA, TEM	Inclusion for healthy controls: siblings, friends or family friends of the children with PCD with no history of chronic or recent acute respiratory problems, no use of medications, and normal physical examination and spirometry.
Pifferi *et al* (2010) (307)	Italy	Dec 2007 to May 2008	Cross-sectional (questionnaires)	78	N/A	Mean (SD): 21.4 (12.9), range 1.7 to 48.5	HSVA + TEM	Not reported

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Pifferi *et al* (2012) (228)	Italy	March 2008 to May 2010	Cross-sectional	50	N/A	Median (IQR) for children: 11 (5.25); n=26 Median (IQR) for adults: 30.5 (9.5), range 18 to 47; n=24	LM, TEM, cell culture	Not reported
Pifferi *et al* (2015) (214)	Italy	Not reported	Cross-sectional	45	53	Median (IQR): 14 (22.25)	nNO, (HSVA + TEM, n=37), (HSVA + cell culture, n=8)	Not reported
Ratjen *et al* (2016) (126)	Canada	Not reported	Cross-sectional +	35	17	Median (IQR): 11.0 (6.8 to 15.3)	Clinical phenotype, nNO,	Inclusion: at least 6 years at enrolment; ability to perform reproducible

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
			prospective cohort study				TEM, genetic testing	spirometry meeting ATS standards; ability to produce sputum spontaneously; clinically stable at the time of assessment Exclusion: use of IV antibiotics or oral quinolones within previous 14 days; use of inhaled antibiotics within the previous 28 days; recent history of pneumothorax or haemoptysis; patients with *P. aeruginosa* or *Burkholderia cepacia*

Appendix C

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								complex infection (for CF only)
Rogers *et al* (2013) (223)	UK	July 2012 to February 2013	Cross-sectional	24	N/A	Median: 15, range 4 to 73	According to international diagnostic guidelines (no further details)	Not reported
Santamaria *et al* (2008) (227)	Italy and the Netherlands	Not reported	Cross-sectional, mixed retrospective and prospective study	20	CF (50) from a previously published cohort of 119 CF patients	Median: 14.3, range 4.6 to 27.5	LM, TEM	Not reported
Santamaria *et al* (2014) (308)	Italy	Not reported	Cross-sectional,	16	42	Median: 10.4, range 4.9 to 17.2	HSVA, TEM	Inclusion: lung disease stability, ability to perform reliable

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
			prospective study					pulmonary function tests, availability of a chest HRCT obtained in stable conditions in the preceding 3 months Exclusion: airway infections and asthma exacerbation 4 weeks before the enrolment; symptomatic heart disease; need for chronic oxygen administration; corticosteroids or bronchodilators use during the previous 2 weeks or 24h, respectively; use of anticonvulsant or

Authors (year of publication)	Countries of data collection	Data collection period	Study design	*n* PCD patients	Reference group (*n*)	Age PCD patients (years)	PCD diagnostics	Inclusion/exclusion criteria
								psychoactive drugs; craniofacial abnormalities, neuromuscular disorders or concomitant genetic diseases such as Trisomy 21 or Prader–Willi syndrome
Shah *et al* (2016) (309)	UK	1980 to 2014	Retrospective cohort study	151	N/A	Median (IQR) in 2014: 35 (26 to 47), range 19 to 75	Clinical phenotype, nNO, LM, TEM; 3% were diagnosed on clinical symptoms alone	Not reported
Shoemark *et al* (2009) (212)	UK	March 2005 to March 2007 and	Case-control	20	Non-PCD bronchiectasis	40 (95%CI 32-45)	LM, TEM	Inclusion for healthy controls: no history of respiratory disease and

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
		January 2006 to June 2006			(20), healthy controls (20)			free from bacterial or viral infections for 8 weeks before study Exclusion: Current smokers, CF patients (screened by sweat test, followed by CF genotyping), history of asthma
Smit *et al* (1996) (310)	The Netherlands	1952 to 1994	Retrospective cohort study	21	N/A	Age at present, (range): 46 (32-61) for lung resection group (n=13); 46	Clinical phenotype (n=8) or TEM + HSVA (n=13)	Exclusion: language barrier, psychiatric problems, and living abroad

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
						(24-66) for group without lung resection (n=8)		
Sunther *et al* (2016) (125)	UK	January 2003 to April 2013	Retrospective cohort study	30	N/A	Median: 11.4, range 6 to 16.2	Clinical phenotype, nNO, HSVA, TEM	Inclusion: aged 6 to 16 years, able to perform spirometry, history of at least one pulmonary exacerbation Exclusion: incomplete set of spirometric assessments
Svobodova *et al* (2013) (311)	Czech Republic	Not reported	Retrospective cohort study	29	N/A	Median: 14.5, range 1.5 to 24	Clinical phenotype, HSVA, TEM,	Not reported

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
							genetic testing (for ODA only)	
Tamalet *et al* (2001) (312)	France	1989 to 1999	Prospective cohort (unclear)	43	N/A	Mean (SD): 5.8 (3.3), range 1 to 13	CBF, TEM	Exclusion: any known pathologic conditions, such as cystic fibrosis, α1-antitrypsin deficiency or immunodeficiency
Valerio *et al* (2012) (313)	Italy	June 2007 to December 2008	Cross sectional study	10	8	Mean (SD): 13.2 (2.8)	LM, TEM	Exclusion: unable to perform spirometry or maximal cardiopulmonary exercise testing, acute upper or lower airway infections, and any concurrent medical illness at the time of the study

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Vallet *et al* (2013) (314)	France	Not reported	Retrospective study	60	N/A	Range 0 to 15	Clinical phenotype, HSVA, TEM	Not reported
Wells *et al* (2011) (315)	Canada	Not reported	Observational study	10	CF (20), healthy controls (20)	Mean (SD): 13.8 (2.3)	Not reported	Inclusion: clinical stability, FEV_1 > 70% predicted, good nutritional status (BMI z-score -2 ± 2) Inclusion for CF: free of a recent pulmonary exacerbation in the 3 months preceding recruitment, normal oral glucose tolerance tests near the time of the magnetic resonance spectroscopy testing

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
Whalley *et al* (2006) (316)	UK	July 2005 to January 2006	Prospective qualitative interview study (matched-pairs design)	12	N/A	Mean: 49.8, range 27 to 65	Not reported	Inclusion: living within 250 km from London Exclusion: < 18 years
Yiallouros *et al* (2015) (317)	Cyprus	1998 to 2013	Cross-sectional	30	N/A	Median: 24.3, range 0.7 to 63.7	TEM + (nNO, HSVA)	Not reported
Zihlif *et al* (2005) (318)	UK	Not reported	Cross sectional prospective study	20	10	Median (IQR): 10.8 (9 to 14)	Clinical phenotype, nNO, CBF, TEM	Inclusion: at least 7 years old, able to perform reproducible spirometry, and stable pulmonary disease Exclusion: positive sputum culture or

Appendix C

Authors (year of publication)	**Countries of data collection**	**Data collection period**	**Study design**	***n* PCD patients**	**Reference group (*n*)**	**Age PCD patients (years)**	**PCD diagnostics**	**Inclusion/exclusion criteria**
								baseline FEV_1 of less than 40% predicted
Zihlif *et al* (2006) (319**)**	UK	Not reported	Cross sectional prospective study	23	11	Median (IQR): 10.3 (9 to 14)	Clinical phenotype, nNO, CBF, TEM	Inclusion: at least 7 years old, able to perform reproducible spirometry, and stable pulmonary disease Exclusion: positive sputum culture or baseline FEV_1 of less than 40% predicted

PCD: Primary ciliary dyskinesia, N/A: not applicable, SD: standard deviation, nNO: nasal nitric oxide, HSVA: high-speed video microscopy analysis, TEM: transmission electron microscopy, IF: immunofluorescence, MBW: multiple breath washout, CF: cystic fibrosis, CBF: ciliary beat frequency, IQ: intelligence quotient, HRCT: high-resolution computed tomography, CF-PI: cystic fibrosis with pulmonary insufficiency, CF-PS: cystic fibrosis with pulmonary sufficiency, LM: light microscopy, CT: computed tomography, IDA: inner dynein arm defect, MRI: magnetic resonance imaging, ATS: American Thoracic Society, ODA: outer dynein arm defect, FEV_1: forced expiratory volume in 1 second, BMI: body mass index.

C.2 Outcome measures definitions

Table 22. Definition of outcome measures included in the clinical outcome measure scoping review, stratified by study outcome and population descriptor.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Abitbul *et al* (2016) (259)	None	CT (bronchiectasis), fertility, microbiology (sputum cultures), spirometry (FEV_1 % predicted)	Not described
Ahmad *et al* (2015) (237)	MBW (correctly categorised %, mean time saved in seconds, mean time saved %, coefficient of variance)	None	MBW: conducted according to published standardised protocol (320). Correctly categorised was defined as % of correctly predicted values using the upper limit of normal, calculated from healthy controls. Reference was 'LCI standard', to which $LCI_{0.75}$, $LCI_{0.5}$ and $LCI_{0.25}$ were compared. Coefficient of variance: calculated from the mean of the coefficient of variance of the intra-test FRC and LCI (SD/mean). Time saved in each of the shortened MBWs is to their respective end-points.

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Alanin *et al* (2015) (278)	Microbiology (period prevalence rate (PePR), period prevalence rate for chronic infection (PePRchr))	Spirometry (FEV_1 and FVC % predicted)	Microbiology: PePR was defined as the percentage of patients who grew the pathogen during a calendar year and PePRchr the percentage of patients who could be classified as chronically infected during a calendar year according to the study criteria detailed below. Criteria and definitions were based on the modified 'CF Leeds criteria' (321). Lung infection status was based on at least 4 samples from the lower airways collected during a period of 1 year and was defined as: a) Chronic infection, when >50% of the preceding 12 months' cultures were positive for the specific pathogen; b) Intermittent colonization, when 50% or less of the preceding 12 months' cultures were positive for the specific pathogen; c) Free of colonization and infection, when no growth has occurred in the lungs in the previous 12 months. However, patients with 2 or 3 positive bacteriological samples in combination with abnormal precipitins were classified as chronically infected.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Behan *et al* (2017) (225)	HR-QoL (QOL-PCD questionnaire, SF-36, shortened SGRQ-C, SNOT-20)	Microbiology (infection with *Pseudomonas aeruginosa*), spirometry (FEV_1 % predicted)	The analyses assessed the extent to which items correlated with their hypothesised versus competing scales; item-to-scale correlations should be ≥0.40 with the intended scale and lower correlations with competing scales. Correlations between 0.50 and 1.00 were interpreted as strong, correlations between 0.30 and 0.50 as moderate, correlations between 0.10 and 0.30 as small and correlations <0.1 as weak, following Cohen's guidelines. 1. The QOL-PCD questionnaire was developed specifically for PCD and consists of 49 items, with most responses captured using a 4-point Likert scale (224, 225, 322, 323). 2. SF-36 was derived from an observational study that began in 1986 on subjects with cardiac impairment. It is a 36-item self-administered questionnaire that includes eight scales, four of which relate to physical health: physical functioning, physical role limitation, bodily pain and general health perception. The remaining four scales are related to mental health: emotional role limitation, mental health, social functioning and vitality. Each scale is scored from 0-100. These eight scales provide two component summary scores: mental component summary and physical component summary in which normal score is 50±10. 3. The SNOT-20 is a validated disease-specific HR-QoL measure for rhinosinusitis

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			that consists of 20 items. Each item is measured on an ordinal Likert scale from 0 to 5, with higher scores indicating worse symptoms. The first 12 items pertain to specific physical sinonasal symptoms including nasal symptoms and ear symptoms. The final 10 items address more systemic and psychological symptoms. 4. SGRQ-C is a disease-specific instrument designed to measure impact on overall health, daily life, and perceived well-being in patients with obstructive airways disease. The shorter 40-item version of the SGRQ does not specify a recall period and has been validated specifically for COPD patients.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Boon *et al* (2014) (279)	None	Anthropometry (weight, height and BMI z-scores), spirometry (FEV_1 and FVC z-score), microbiology (life-time prevalence), chest radiographs and CT	Chest radiographs or CT scans: presence or absence of pulmonary infiltrates, lobar consolidation/atelectasis and bronchiectasis. Microbiology: Sputum, bronchoalveolar lavage or cough swabs available since diagnosis were evaluated for the presence of respiratory pathogens, and lifetime prevalence was reported as 'has ever had infection with'. Chronic colonisation by pathogen was defined as persistence of the same bacteria in at least 3 sputum samples over a period of at least 6 months. Anthropometry: weight, height and BMI were reported as z-scores, according to Flemish growth curves (324). Spirometry: z-scores reported according to Quanjer equations (191).

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Boon *et al* (2015) (280)	Spirometry (FEV_1, FVC, FEV_1/FVC and $FEF_{25–75}$ z-scores), N_2 MBW (LCI), HRCT	Anthropometry (weight, height and BMI z-scores)	MBW: LCI was calculated by dividing the cumulative expired volume by the functional residual volume. At least two technically acceptable measurements per patient were performed. S_{cond} and S_{acin} were both multiplied by tidal volume to normalise for age, as proposed in the MBW consensus guidelines (325). The mean LCI of at least two technically acceptable measurements was used. Spirometry was performed according to the ATS/ERS guidelines (326). FEV_1, FVC, FEV_1/FVC and $FEF_{25–75}$ were expressed as z-scores according to the reference equations from the GLI. A z-score below −1.96 was defined as abnormal. Spirometry was performed on the same day as MBW. HRCT: A cystic fibrosis computed tomography (CFCT) score, a variant of the modified Brody Score, was used to quantify specific abnormalities on chest CT: severity and extent of bronchiectasis, severity and extent of airway wall thickening, mucus plugging in central and peripheral airways, parenchymal abnormalities (consolidation, atelectasis, cysts and ground glass opacities) and air trapping. The lingula was considered as a separate lobe. Scores were expressed as percentage of the maximum score of 207 and a total CFCT score >5% was defined as abnormal. Anthropometry: height, weight and BMI were expressed as z-scores according to Flemish reference equations (324).

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Bush *et al* (2006) (281)	Inflammatory markers (IL-8 concentration), sputum biophysical and transport properties (dynamic viscoelasticity, wettability, cohesivity, interfacial tension, solids composition, DNA, IL-8 concentration, cough transportability)	Spirometry (FEV_1 and FVC % predicted), microbiology (chronic infection with *Pseudomonas aeruginosa*)	Spirometry: performed according to ATS guidelines (326). Three reproducibility flow-volume curves with <10% variability in FEV_1 were recorded. Sputum properties: viscosity was defined as the loss of energy from a rheologic probe (stress) and thus the resistance to flow. Elasticity referred to the recoil energy transmitted back to the probe. Cohesivity was defined as interfacial tension multiplied by the new area as after a test substance is subjected to nonshearing stress. Interfacial tension measured the interfacial tension at the sputum/air interface. Sputum was collected during exacerbation, which was defined only by the centre physician's decision to begin antibiotic therapy at clinic visit.

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Carotenuto *et al* (2013) (282)	HR-QoL (Wechsler Intelligence Scale for Children-III edition (WISC-III), Child Behavior CheckList (CBCL) questionnaire, Parental stress index-short form (PSI/SF))	Anthropometry (BMI)	Intelligence assessment: WISC-III is composed of 13 distinct subtests with 6 verbal scales including language-based items, whereas the 7 performance scales consist of visual-motor items that are less dependent on language. 5 of the subsets in each scale produce scale-specific IQs as verbal IQ and performance IQ and the 10 subtest scores produce a total scale IQ. Behavioural assessment from CBCL: mothers were instructed to answer questions about their child's behaviour during the past 6 months. Items are scored as 0=not true, 1=somewhat true or sometimes true, or 2=very true or often true. The questionnaire yields 8 factors: withdrawn, somatic complaints, anxious/depressed, social problems, thought problems, attention-hyperactive, rule-breaking behaviour, and aggressive behaviour; as well as 3 global scores for externalizing and internalizing behaviours and total behaviour score. PSI/SF: yields scores of maternal stress across 4 domains: parental distress, parent-child dysfunctional interaction, difficult child, and total stress. Each item was graded on a 5-point Likert scale, with higher scores indicated higher perceived stress in the parents. A score at, or above, the 85th percentile indicates high stress level.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Cockx *et al* (2017) (283)	Inflammatory markers (Chemotactic response of PCD neutrophils to 4 chemoattractant: C5a, LTB4, chemokine CXCL5 and chemokine CXCL8)	Spirometry (FEV_1 and FVC % predicted), microbiology	Migration of the PCD polymorphonuclear neutrophils was expressed relative to migration of the reference adult control.
Cohen-Cymberknoh *et al* (2017) (284)	Microbiology (colonized vs non-colonized with Pseudomonas aeruginosa (PA)), spirometry (FEV_1 % predicted), CT	Anthropometry (BMI percentile for ≤20 years and BMI for >20 years)	Microbiology: Several definitions of colonized and non-colonized with PA were used. Only a few patients in the study could meet the Leeds criteria (321), which is the most rigorous criteria and used in CF. Therefore, patients were classified as non-colonized if they had never been cultured with PA or cultured only once whereas colonized patients were defined as having had least two positive sputum cultures for PA during the study period. Colonized groups were defined as having a) at least 4 positive cultures during the study period (n = 41), b) at least 6 positive cultures during the study period (n = 28) or c) two or more consecutive positive cultures or two consecutive years with at least one positive PA culture each year(n = 54).

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			Spirometry: Decline of FEV_1% predicted throughout the study period was calculated numerically by subtracting the first best FEV_1 from the last one in the study, divided by the number of years each participant took part in the study. CT: Brody scores were calculated with a slight modification (hyperaeration of the lungs was assessed instead of air trapping).

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Cohen-Cymberknoh *et al* (2014) (105)	HRCT, spirometry (FEV_1 % predicted), microbiology	Anthropometry (BMI percentile)	Pancreatic insufficiency was defined as stool elastase <100µg/g stool or coefficient of fat absorption < 93%. Spirometry: pulmonary function tests were performed according to ATS/ERS guidelines (326). FEV_1 was presented as % predicted, according to Wang *et al* (327) for children and Hankinson *et al* (328) for adults. HRCT: each lung lobe, including the lingula, was counted as a separate lobe. The Brody score was calculated with a slight modification: hyperaeration of the lungs was evaluated instead of air trapping, as expiratory images were not obtained in all patients. Sub-scores for the presence and severity of bronchiectasis, mucous plugging, bronchial wall thickening, parenchyma, and focal hyperaeration in each lobe were calculated. Parenchymal findings of ground glass, consolidation, and cysts or bullae were all considered in determining a single parenchyma sub-score. The sum of sub-scores constituted lung total Brody scores for each patient. Microbiology: chronic infection was defined when patients had at least three positive sputum cultures within 1 year.

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Davis *et al* (2015) (285)	Anthropometry (height, weight and BMI percentile), spirometry (FEV_1 and FEF_{25-75} % predicted, infant $FEV_{0.5}$, infant $_{25-75}$ z-score), CT (n lobes with bronchiectasis, n lobes with alveolar consolidation)	Microbiology	Spirometry: performed according to ATS/ERS criteria and overread for quality (326). Spirometric measurements were expressed as percent predicted (329) and infant lung function as z-scores (330, 331). Chest CT images were scored for the presence of bronchiectasis and parenchymal disease in six lobes, including the lingula as a lobe, using the Brody score.
Ellerman *et al* (1997) (124)	Spirometry (FEV_1 and FVC % predicted)	Chest radiography (presence of bronchiectasis), microbiology	Spirometry: the best of 3 valid attempts was used as outcome. Published reference values for children (332) and complied Danish reference values for adults were used. Spirometry was measured 3 to 4 times per year and the annual lung function is reported as the mean of the measurements performed at the clinic during the previous years.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Frija-Masson *et al* (2017) (122)	None	Spirometry (FEV_1 % predicted, FEV_1, FVC, FVC, TLC, TLC, FEV_1/FVC % predicted), microbiology (with and without chronic *Pseudomonas aeruginosa* infection), HRCT (modified Bhalla score), dyspnoea score (Modified Medical Research Council scale), treatment (number of courses of antibiotics (IV, oral, inhaled)), fertility, lobectomy (long-term oxygen use, lung transplant), mortality	Spirometry: performed according to the ERS/ATS guidelines (326). Postbronchodilator FEV_1 was used and FEV_1 decline was calculated if there were 3 or more values of FEV_1 and a follow-up of at least 2 years. Annual decline was calculated according to the European Coal and Steel Community (ECSC)/ERS 1993 reference equation (333). Microbiology: chronic infection was defined as those with a positive pathogen in at least 3 sputum samples in less than 6 months. CT scoring system: modified Bhalla score for chest bronchiectasis. In patients with situs inversus, the lung in which the middle lobe was identified was considered as the right lung. The scores from the 6 lobes were summed to provide a total score ranging from 0 (normal) to 48 (maximal score).

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Gokdemir *et al* (2014) (232)	Spirometry (FEV_1, FVC, PEF, FEF_{25-75} % predicted), comfort and efficacy (questionnaire score), SpO_2	Anthropometry (weight and height z-scores)	Spirometry: performed according to the ERS/ATS guidelines (326). Measurements were taken at the same time of the day before and after 30 min period following the last treatment session of conventional pulmonary rehabilitation or high-frequency chest wall oscillation on the 1st and 5th day. SpO_2 was measured transcutaneously at rest, for 5 min immediately before, 30 min during and 30 min immediately following each session. SpO2 was measured with a fingertip pulse oximeter. Perceived efficiency and comfort level: patients completed a written questionnaire to rate comfort and efficiency of the two modalities with a 5-point scale (extremely = 4, very = 3, somewhat = 2, not very = 1, and not at all = 0).
Green *et al* (2012) (112)	SF_6 MBW (LCI absolute values and z-scores, LCI within-session variability (CV)%, S_{acin}, S_{cond}, FRC_{SF6}), spirometry (FEV_1, FVC,	Anthropometry, microbiology	MBW: LCI was calculated as the number of lung volume turnovers (the cumulative expired volume divided by the functional residual capacity) needed to lower the end-tidal tracer gas concentration to less than 1/40th of the starting concentration. The mean LCI result from 3 MBW measurements in each patient was used for analysis. Spirometry: performed according to ATS/ERS standards (326, 329). Abnormal lung function was defined as z-scores < -1.96. The upper limit of normal was defined as

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	FEF_{25-75}, FEV_1/FVC ratio z-scores)		the predicted mean plus 1.96 SD for MBW variables and the lower limit of normal as predicted minus 1.96 SD for spirometry variables. Spirometry parameters were calculated using the British growth reference charts. MBW z-scores calculated using Swedish normative data.
Green *et al* (2016) (231)	MBW ($LCI_{2.5}$, $LCI_{3.0}$, $LCI_{4.0}$, $LCI_{5.0}$, $LCI_{7.0}$, $LCI_{9.0}$), spirometry (FEV_1, FVC, FEV_1/FVC ratio and $MMEF_{25-75}$ z-scores)	Anthropometry (weight, height, BMI z-scores)	Spirometry was performed according to ATS/ERS guidelines (326). GLI reference equation was used to obtain z-scores and a z-score < -1.64 was considered an abnormal spirometric value.
Hellinckx *et al* (1998) (286)	Spirometry (FEV_1, FVC, change in FEV_1 and FVC % predicted), body plethysmography (thoracic gas volume, total lung capacity,	None	Spirometry: according to ERS guidelines, the best of 3 maximal expiratory flow volume manoeuvres was analysed (326). All measurements were expressed as % of predicted values for sex and height according to Zapletal *et al* (334). Body plethysmography: single breath diffusing capacity and Krogh factor were measures according to ERS guidelines (335). FEV_1, vital capacity and Raw % predicted were calculated according to Zapletal *et al* (334). TLC, RV, thoracic gas volume and single breath diffusing capacity % predicted were calculated according

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	residual volume, and airway resistance)		to ERS guidelines. Reference values for total respiratory system resistance and reactance were according to Duiverman *et al* (336). All tests were done before and 20 min after administration of 200 µg of salbutamol. Drug dose was chosen according to Bibi *et al* (337).
Irving *et al* (2013) (118)	Spirometry (FEV_1, FVC and MEF_{25-75} z-scores), MBW (LCI and functional residual capacity), HRCT	Microbiology (infection with *Pseudomonas aeruginosa*)	Spirometry: performed according to ATS/ERS recommendations (326). Subjects completed a minimum of 3 forced expiratory manoeuvres, and FEV_1 (L) and FVC (L) were expressed as z-scores. MBW: LCI was defined as the number of volume turnovers of the lungs required to reduce an inert gas to $1/40^{th}$ of its starting concentration. Minimum of 2 of the 3 tests had to meet the acceptability criteria to be included in the analyses. HRCT: presence and severity of specific CT features was recorded for each lobe (individual scoring system), including extent of bronchiectasis, severity of bronchiectasis, bronchial wall thickness, small and large mucus plugs, and air trapping. Used a study-specific score that was then compared to the Brody score. Chronic infection with *Pseudomonas aeruginosa* was defined as at least 2 positive cultures on cough swab or sputum culture over the last 5 years.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Irving *et al* (2017) (230)	Spirometry (FEV_1 z-score), MBW (LCI)	None	MBW: LCI was calculated as the mean of at least 2 acceptable tests. Spirometry: performed according to ATS/ERS guidelines (326).
Jain *et al* (2007) (216)	Chest radiography (dextrocardia, hyperinflation, bronchial wall thickening and dilation, mottled shadows, consolidation or collapse), HRCT	Microbiology	Chest radiography: modified Chrispin-Norman score (no need for lateral film) (338). Lungs were divided into 4 zones on the frontal film: right upper, left upper, right lower, left lower; the following were scored for each zone: bronchial wall thickening, ring shadows, mottled shadows, and large soft-tissue shadows; scores of 0 (not present), 1 (present but not marked), and 2 (marked) were given for each of these 4 parenchymal lung features. Radiographs were also assessed for over-inflation, with a possible maximum score of 6. HRCT: Brody score used to evaluate 5 features independently in each lobe (bronchiectasis, mucus plugging, peribronchial thickening, parenchymal changes of consolidation and ground-glass density, and focal air-trapping).

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Joensen *et al* (2014) (216)	Breath profiles (volatile organic compounds), microbiology (chronic infection), number of exacerbations	Spirometry (FEV_1 and FVC % predicted)	Microbiology: chronic infection was defined by the Copenhagen criteria (persistent presence of pathogen in microbiological culture samples for at least 6 consecutive months, or less when combined with the presence of 2 or more *Pseudomonas aeruginosa* precipitins). Samples were obtained by expectoration sputum, endo-laryngeal suctioning and bronchoalveolar lavage. Pulmonary exacerbation was defined as need to start additional antibiotic therapy and presence of at least 2 of the following criteria: change in sputum volume and/or colour, increased coughing, increased lethargy, feeling unwell, or increased need for sleep, decreased appetite or weight loss, decrease in lung function ≥10%, increased shortness of breath or new acquired radiological changes. Spirometry: performed according to the ATS/ERS guidelines (326). Exhaled breath sampling: 2 measurements per patient were performed with an interval of 5 minutes between them.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Kawakami *et al* (1996) (288)	Chronic sputum production (duration throughout the year, daily amount, colour), sputum and nasal scores	Fertility (sperm motility)	Sputum and nasal scores were calculated to estimate the severity of the symptoms using the answer from the patients in the following manner. Most severe symptoms for each question were valued at 30. Scores were obtained by summing the points from the five questions concerning chronic sputum production and from the six questions concerning chronic nasal symptoms, respectively. The maximum possible scores for the sputum and the nose were 150 and 180 respectively and 0 indicated that they had no symptoms. Chronic sputum production: obtained from questionnaires sent to patients. Questions included duration of sputum production throughout the year, daily amount of sputum and colour of mucus.
Kennedy *et al* (2007a) (289)	Lythoptysis (symptoms), radiographic findings (calcium deposition)	Spirometry (FEV_1 % predicted), microbiology, lobectomy	Spirometry: FEV_1 used was the best +/- 1 year of when the CT scan was performed. Symptoms of lithoptysis: spitting up a hard concretion, a firm stone-like structure in the sputum or a gritty sensation in the sputum. Radiographic findings: evidence of calcification.
Kennedy *et al* (2007b) (119)	HRCT (study-specific score)	Spirometry (FEV_1 % predicted), microbiology, lobectomy	High-resolution CT images were assessed for severity of bronchiectasis in each lobe. A score of 0 indicated no bronchiectasis; 1, mild bronchiectasis (bronchial dilatation 2 times the diameter of the accompanying blood vessel); 2, moderate

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			bronchiectasis (bronchial dilatation 2 to 3 times vessel diameter); 3, severe bronchiectasis (bronchial dilatation more than 3 times vessels diameter). An overall bronchiectasis severity score for all 6 lobes was calculated (score range 0-18). The distribution of bronchiectasis was classified in each lobe as central (proximal 50% of lung parenchyma), or diffuse. If lobectomy was performed, a severity score of 3 was assigned to the missing lobe by arbitrary definition, and distribution was presumed diffuse. The presence or absence of peribronchial thickening and mucous plugging for each lobe was recorded. Other radiographic findings included: mucous plugging, peribronchial consolidation, lobar collapse and atelectasis, pleural effusion, nonspecific infiltrate, emphysema, calcium deposition, pectus excavatum)
Knowles *et al* (2014) (339)	None	Spirometry (FEV_1 % predicted), fertility (status)	Spirometry: FEV_1 % predicted was calculated using ERS Task Force multi-ethnic reference values. The latest available FEV_1 was used for the *RSPH1* individuals and the value recorded at the research visit for the classic PCD cases.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Koh *et al* (2000) (108)	Spirometry (FEV_1 % predicted, $\Delta FFEV_1$, PC_{20} (provocation concentration of metacholine producing a 20% fall in FEV_1, $M\Delta FFEV_1$)	None	Spirometry was performed after 3 weeks of regular use of medication. The largest value of the triplicate FEV_1 at each time point was adopted for analysis. High-dose methacholine inhalation tests were carried out by using a modification of the method described by Chai *et al* (340). Each subject inhaled 5 inspiratory capacity breaths of buffered saline solution and increasing concentrations of methacholine at 5-min intervals. FEV_1 was measured 60 to 90 s after inhalation of each concentration level. The procedure was terminated when FEV_1 had fallen by >40% from the post-saline value, or when a maximal response plateau had been established. This was considered to occur if 3 or more data points of the highest concentration fell within a 5% response range. An additional 5 or 10 inhalations of the 200 ug/mL solution were taken if the last three data points of less than a 40% fall did not satisfy the above criteria. The response was expressed as the % fall in FEV_1 ($\Delta FFEV_1$) from the post-saline solution value and was plotted against logged concentrations of inhaled methacholine. The dose-response curves were characterised by their position and maximal response. The position was expressed as PCD_{20}, which was calculated by log-linear interpolation between 2 adjacent data points. The maximal airway response plateau ($M\Delta FFEV_1$) was defined as the level of

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			maximal response plateau by averaging the consecutive points on the plateau. The last data point of the dose-response curve was used if a plateau could not be achieved.
Li *et al* (2005) (290)	HRCT (distribution of bronchiectasis)	Spirometry (FEV_1 and FVC % predicted), microbiology	HRCT: presence or absence of bronchiectasis was recorded in each lobe, with the lingula being considered as a separate lobe. Widespread disease was defined as bronchiectasis involvement of 5 or more lobes. Spirometry: performed according to the ATS guidelines (326). Three technically acceptable manoeuvres were performed each time, and the highest value of FEV_1 and its corresponding FVC were recorded. Bronchiectasis was defined as idiopathic if extensive investigations failed to reveal an underlying aetiology. The commonest organism isolated for each aetiology were reported.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Loomba *et al* (2017) (291)	Spirometry (FEV_1, FVC, FEV_1/FVC and FEF_{25-75} % predicted), exercise testing (peak VO_2 absolute values and % predicted, peak $EtCO_2$, exercise time, resting O_2 saturation, % increase in blood pressure, arrythmia during exercise test)	None	Exercise testing: modified Bruce protocol. Those undergoing cardiopulmonary exercise testing using a cycle ergometer, there was a warm-up period followed by a progressive exercise test with a modified Godfrey protocol (341). Ventilatory data were obtained every 15 seconds.

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Lopes *et al* (2015) (215)	Spirometry & body plethysmography (FVC, FEV_1, FEV_1/FVC, PEF, $FEF_{25-75\%}$, TLC, RV and RV/TLC % predicted, DLco % predicted, % bronchodilator response), HRCT, dyspnoea	Anthropometry (BMI), treatment (use of inhaled medication (bronchodilator, corticosteroids, antibiotics, DNase))	Dyspnoea: modified Medical Research Council (MRC) scale (342). Spirometry/body plethysmography: All tests followed the standards formulated by the ATS (326). Bronchodilator response was identified based on the presence of a variation of 12% and 200 mL in FEV_1 or FVC after the use of 400ug of inhaled salbutamol. Pereira's and Neder's equations were used in the interpretation of the functional parameters (343, 344). Airflow obstruction was defined by an FEV_1/FVC value <70% predicted. A restrictive pattern was defined as the presence of a TLC <80% of predicted; this cut off point was also used to define abnormality in DLco. HRCT: extent of bronchiectasis was established by the modified scale described by Bhalla *et al* (219), which ranges from 0 to 18. Each lung lobe (considering the lingual and middle lobes as independent) was scores as follows: 0 = no bronchiectasis; 1= one or partial bronchopulmonary segment involved; 2 = two or more bronchopulmonary segments involved; and 3 = generalized cystic bronchiectasis.

Madsen *et al* (2013) (292)	N_2 MBW (LCI, S_{cond}, S_{acin}, FRC_{N2}), spirometry (FEV_1, FVC, FEV_1/FVC, FEF_{25-75} and TLC z-scores), body plethysmography (sRaw, FRC, RV, TLC, VC, RV/TLC z-score, Dlco and Dlco/V_A), exercise testing (VO_{2peak} absolute value, % predicted and z-score, maximal heart rate, test duration, oxygen pulse, maximum workload corrected for body weight, FR, VT, RER, VR, VE, $V_E/VCO_{2,\ anaerobic\ threshold\ \%\ predcited}$), HR-	Anthropometry (BMI z-scores), microbiology	VO_{2peak}: a valid peak was defined by continuous objective signs of exhaustion during verbal encouragement from the test leader, combined with at least one of the following criteria: respiratory exchange ratio >1 at test termination, or maximal heart rate > 85% of age-based predicted maximum. The VR reflecting ventilatory capacity was calculated, as was the ventilatory equivalent of CO_2 (V_E/VCO_2) reflecting efficacy of ventilation. VR< 15% or V_E/VCO_2>40 were considered abnormal and to be positive signs of ventilatory limitation during the test. Reference values of VO_{2peak} were derived from comparable assessment in 937 healthy Danish children and young adults and this reference material was evaluated and compared with the group of matched healthy controls. Spirometry & body plethysmography: all-ages reference equations were used (329). For children, the reference equations of Koopman *et al* (345) were used for DLco and Zapetal *et al* (334) for whole-body plethysmography, except sRaw for which the reference equation of Kirby *et al* (346) was used. For adults (>18 years), reference equation of Cotes *et al* and Quanjer *et al* (347, 348) for DLCO and whole-body plethysmography were used, respectively. N_2 MBW: Calculated LCI and the normalized phase III slope indices S_{cond} and S_{acin} using pre-reviewed normative data as reference material. HR-QoL: selected and combined validated questions from the SGRQ, CF Questionnaire (CFQ-R), SNOT-22 and SF-36, to extract simple questions about physical activity and limitations that were useful for the study. All, including healthy

	QoL (study-specific questionnaire)		control subjects, answered questions on the following subjects: physical limitations in activities of every-day-life due to symptoms, subjective judgement of the difficulty performing vigorous activities, and weekly hours spent on physical activities, such as running, cycling and sports. Abnormal lung function and VO_{2peak} was defined as z-score <-1.96, whereas abnormal LCI was defined as z-score >1.96. Chronic PSA: chronic infection with P aeruginosa, defined as more than 50% of positive airway cultures the previous year. Intermittent *P aeruginosa*: intermittent infection with *P aeruginosa*, defined as least one positive culture in the last year. Chronic XA: chronic infection with *Achromobacter xylosoxidans*, defined as more than 50% of positive airway cultures the previous year.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Maglione *et al* (2012) (293)	Spirometry (FEV_1, FVC, FEV_1/FVC and FEF_{25-75} z-scores, change in FEV_1 z-score), HRCT	None	Definition of stability: partly modified definition of stability previously suggested in CF (349). Stable patients were those with no recent change (preceding 4 weeks) in chest physical examination, sputum volume or colour, dyspnoea, cough frequency, malaise, fatigue, or weight. Definition of unstable lung disease: febrile, illness indicating substantial infectious insult, and/or worsening symptoms suggesting progression of bronchiectasis, that were unresponsive to prolonged oral and/or IV Abx and daily physiotherapy with nebulized saline. In the absence of any generally agreed protocol or evidence, the decision to perform a second CT scan was also made on an individual basis after discussion with the patient and his family. HRCT scan scoring: modified Brody scoring system (227). Bronchiectasis score range 0-12, mucus plugging score (range 0 to 6), peribronchial thickening score (0 to 9), parenchyma score (0 to 9), mosaic perfusion score (0 to 4.5). A score was calculated for each abnormalities, and these scores were summed to provide a total score for each lobe. The scores from the 6 lobes were then summed to provide a total HRCT scan score, with a theoretical range from 0 (normal) to 243 (maximal score in all lobes). In practice the maximal score could not exceed 207, since a lobe cannot have more than 2/3 involvement from all abnormalities at the same time. All scores

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			were normalized to a scale of 0-100, representing a percentage of a maximum possible score, and a total score of >5% was abnormal. Spirometry: measured according to published criteria (326). The best of 3 valid attempts was used in the analysis. FEV_1 z-score <-1.96 was defined as abnormal. Acceptability was checked by an independent blind reviewer inspecting the spirometry loops. The changes in the scores between the 2 evaluations were calculated. A positive value for CT score changes indicated that lung structure abnormalities worsened, while a positive value for change in spirometry indicated an improvement in LF. Spirometry remained stable if the change in FEV_1 % predicted between the 2 evaluations was of no more than +/-10%.
Maglione *et al* (2014a) (208)	Anthropometry (height, weight and BMI z-scores), spirometry (FEV_1, FVC and FEF_{25-75} %	None	Spirometry: according to published criteria (326). FEV_1 z-score <-1.96 was considered abnormal. Anthropometry: BMI z-scores were calculated according to Cole *et al* (350). Microbiology: chronic pseudomonal airway infection: presence of *Pseudomonas aeruginosa* for at least 6 months, with at least 3 positive cultures.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	predicted and z-scores), microbiology		
Maglione *et al* (2014b) (294)	HR-QoL (SGRQ, Leicester Cough Questionnaire, SF-36), spirometry (FEV_1, FVC and FEF_{25-75} % predicted), exercise testing (6-min walk test)	Exacerbations (number of respiratory exacerbations, courses of antibiotics), microbiology (% positive sputum cultures)	Respiratory exacerbation: required systemic antibiotics.

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Maglione *et al* (2017) (295)	MRI, CT	Spirometry (FEV_1, FVC, FEV_1/FVC and $FEF_{25–75}$ z-scores), anthropometry (height, weight and BMI z-scores), treatment (courses of antibiotics, hospital admissions), microbiology (sputum)	Pancreatic insufficiency: stool elastase <100 µg/g. Spirometry: FEV_1 z-score < −1.64 was considered abnormal. Chronic airway infection: same pathogen was detected, after adequate antibiotic therapy, in at least three consecutive cultures within 6 months. HRCT and MRI: morphologic scoring system, originally developed for CF by Helbich *et al* (220), later modified by Puderbach *et al* (351). Maximum achievable total score was 25, indicating the most severe lung changes. For the purpose of quantifying the severity of PCD or CF lung structure deterioration, the total MR score into mild (scores 0-9); moderate (scores 10-18); and severe (scores 19-25). For the categories “severity of bronchiectasis” and “severity of peribronchial wall thickening”, the most prevalent degree of severity was recorded. If mucous plugging was seen within the periphery of a lung segment, bronchiectasis was scored also in that segment. Six lobes were examined, the lingula being scored as a separate lobe. In patients with situs viscerum inversus, the right lung was the lung in which the middle lobar bronchus and the corresponding middle lobe were identified at scans.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Magnin *et al* (2012) (221)	Spirometry (FEV_1, FVC, FEV_1/FVC and FEF_{25-75} z-scores), arterialised capillary blood gases (oxygen (PaO_2) and carbon dioxide ($PaCO_2$) tensions), CT	None	Stability: applied definition accepted in CF (no weight loss or fever, no subjective change in cough frequency, sputum volume and/or colour, and no worsening of dyspnoea). Arterialised capillary blood gases were obtained using a technique described in Gaultier *et al* (352). Spirometry: the best curve out of 2 reproducible expiratory curves were recorded. Beta-agonists were withheld for 12 hours before lung function test, as recommended. Chest CT: protocols varied over time. Chest CT examination protocols have been standardised in accordance with the national recommendations from the French Society of Pediatric Radiology (SFIPP) (i.e. parameters and doses) since 2003. To describe the structural impairment of the lung, items from Bhalla's and Brody's CT scoring systems were used and slightly modified to obtain a score easy to use in routine practice (218, 219). The score described five items (bronchiectasis, mucous plugging, peribronchial thickening, parenchymal abnormalities, and pulmonary hyperinflation), in six pulmonary regions, each lung divided into three regions: (i) the upper region was described from the apex to the tracheal carina, (ii) the middle region from the carina to the lower pulmonary veins, (iii) the lower region from the

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			lower pulmonary veins to the bases. In each region, 0 point was given for absence and 1 point for presence of the following items: mucous plugging, peribronchial thickening, parenchymal abnormalities (condensation and collapse), and pulmonary hyperinflation. Likewise, bronchiectasis were absent (0 point), or present with different degrees of severity assessed by the comparison with the adjacent pulmonary arteria (APA), as proposed in Bhalla's and Brody's CT scoring systems: 1 point for mild bronchiectasis (1–2 times larger than the APA), 2 points for moderate bronchiectasis (2–3 times larger than the APA), and 3 points for severe bronchiectasis (up to 3 times larger than the APA). Additional points were assessed on a CT each time the patient had history of lung surgery: 5 points for lobectomy and 2 points for partial lobectomy. The score range from 0 to 42 points without the points assessed for surgery.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Marthin *et al* (2010) (120)	Spirometry (FEV_1 and FVC % predicted)	HRCT (bronchiectasis), Chest radiography (chronic abnormalities)	Spirometry: for each child every flow–volume curve was evaluated and excluded if technique was insufficient. FEV_1 and FVC measurements were as per ATS standards (326). Longitudinal lung function measurements in each subject following diagnosis were analysed using linear regression on time since diagnosis, for each subject separately, yielding subject-specific estimates of slope. From these slopes, each patient was grouped according to whether the course of lung function increased overall ≥10% points, stabilised (change within 10% points), or decreased ≥10% points in predicted values.
McManus *et al* (2003) (297)	HR-QoL (SGRQ scores on symptoms, activity and impact, SF-36 measures of Health Status physical and mental component scores)	Treatment (use of antibiotics)	Respiratory symptoms were assessed by SGRQ, which provides 3 separate scales (symptoms, activity and impact). The scores are scales in the range 0 to 100, where a score of 100 indicates optimal functioning within the context of respiratory illness. Health Status overall was assessed by version 2 of the SF-36 questionnaire, which is a widely used generic instrument for assessing mental and physical functioning, for which UK population norms are also available. The questionnaire has 8 sub-scales which can be divided into 2 broad groups: physical functioning, role physical, bodily

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			pain and general health, which are primarily physical, and energy/vitality, social functioning, role emotional and mental health, which are primarily mental. The 8 sub-scales are each scored in the range 0 to 100, where a score of 100 indicates optimal functioning. The physical and the mental component scores have well-described population norms. Respondents indicated the extent to which the symptoms had affected them over the past 4 weeks, using 5 categories: 'not at all' (scored 0), 'one day or so' (scored 1), 'a few days a month' (scored 2), 'several days a week' (scored 3), 'almost everyday' (scored 4).
McManus *et al* (2006) (298)	HR-QoL (SGRQ: symptoms, activity, impacts; SF-36 questionnaire: PCS, MCS; General Health Questionnaire; 'Big Five' personality	None	Same as above for SGRQ and SF-36 (297). Stress levels were assessed using the 12-item version of the General Health Questionnaire (GHQ). Each item is on a 4-point scale and the 4 levels on each question are given scores of 0, 1, 2 or 3, with 3 being the most serious. This scale has a range of 0 to 36, and is approximately normally distributed in the population. The 'Big Five' personality dimensions of the Five-Factor Theory were assessed using a modified adjective checklist (353-355). Stigma questionnaire: study-specific measure. Used the stigma sub-scale of the

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	dimensions, stigma questionnaire)		PDQ-39, which is used to assess quality of life in Parkinson's disease, as a model on which to base and develop study-specific questions (356).
Mirra *et al* (2015) (299)	Vitamin D (total 25(OH)D), body plethysmography (FVC, FEV_1, FEF_{25-75}, FRC, RV and FEV_1/FVC % predicted), HR-QoL (SGRQ), physical activity assessment (questionnaire), microbiology	Anthropometry (BMI), HRCT (bronchiectasis), treatment (number of courses of antibiotics)	Vitamin D levels: categorized as being sufficient when >30 ng/ml (>75 nmol/L), insufficient between 20 and 30 ng/ml (50 and 75 nmol/L), and deficient when <20 ng/ml (<50 nmol/L) Self-reported physical activity: assessed using a previously published questionnaire by Madsen *et al* (292). Microbiology: chronic bacterial colonization was defined as persistence of specific bacteria for at least 6 months, with at least 3 positive cultures.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Montella *et al* (2009a) (229)	HRCT, MRI, body plethysmograghy (FEV_1 and FVC % predicted)	Microbiology	MRI and HRCT scores: modified version of Helbich *et al* (220). The severity of mosaic perfusion was excluded as it could not be assessed by morphological MRI. The maximum score was 25 points (instead of the original 27). For the categories "severity of bronchiectasis" and "severity of peribronchial wall thickening", the most prevelant degree of severity was recorded. It was not possible to assess peribronchial wall thickening in the presence of mucous plugging. Hyperintensity on HASTE images had to be present for an MRI diagnosis of mucous plugging. If mucous plugging was seen within the periphery of a lung segment, bronchiectasis was scored also in that segment. Sacculations and abscesses were defined as circular structures with a minimum diameter of 1.5 cm that were air-filled or showed an air-fluid level. A size of 2 cm was required for a diagnosis of collapse and consolidation. Emphysema was defined as an area of decreased signal (compared with the surrounding lung parenchyma) due to a reduction of vessel and parenchymal density. In case of lobectomy or segmentectomy, the maximum scores for "severity of bronchiectasis" and "severity of collapse/consolidation" were arbitrarily assigned to the missing lobe/segments. The assessment of "extent of bronchiectasis" took into account the number of missing segments. Six lobes were examined; the lingula was scored separately. In patients with situs viscerum

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			inversus, the right lung was the lung in which the middle lobar bronchus and the corresponding middle lobe were identified at scans. Body plethysmography: performed according to ATS criteria. FEV_1 > 85% predicted was considered normal.
Montella *et al* (2009b) (300)	HRCT, MRI	None	Same as above for HRCT and MRI (229).
Montuschi *et al* (2014) (301)	Breath profiles (ethanol, methanol, saturated fatty acids, formate, lactate, acetate, leucine/isoleucine, isobutyrate,	Spirometry (FEV_1 and FVC % predicted), microbiology (sputum culture), anthropometry (BMI), treatment (inhaled medication)	Not reported – correspondence, therefore limited information available.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	glutamine/glutamic acide)		
Noone *et al* (1999) (233)	Clearance during cough (mean clearance rates (%/min)), sputum production rate (sputum rheology and ion content (Avg Log G, cough-clearance index, mucociliary-clearance index, Na+ content, Cl- content))	Spirometry (FEV_1 % predicted), cough questionnaire (cough severity and type, amount, ease of expectoration, and nature of sputum, chest tightness, and wheezing)	Studied clearance during a series of controlled coughs from t = 20 to 60 min (t = 0 to 20 min represents the period of delivery of solution). The total number of coughs (spontaneous plus controlled) was limited to 90 during the 60-min period by having each subject cough under the direction of the investigators into a spirometer. Sputum was obtained during the cough manoeuvres as soon as possible after aerosol delivery was completed. Sputum production rate: if a subject produced X grams of sputum Y minutes after the commencement of the study, the sputum production rate was calculated as X/Y grams per minute for that individual. Questionnaire: Before and after aerosol dosing, patients were asked to score, on a questionnaire sheet, the severity and type of their cough, amount, ease of expectoration, and nature of sputum, chest tightness, and wheezing, on a scale of 0 to 10. They were also asked to record comments about any symptoms or feelings in the chest after inhalation.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Noone *et al* (2004) (38)	None	Spirometry (FEV_1 % predicted), microbiology (sputum), radiographs (presence of bronchiectasis), cough (number)	Bronchiectasis was primarily diagnosed clinically based on history of chronic excess mucopurulent sputum production associated with finger clubbing, and, where available, computed tomographic scans of the thorax or with clear abnormalities on chest radiographs were also used to support the diagnosis.
Nyilas *et al* (2017) (304)	MBW/SBW ($LCI_{2.5\%}$, $LCI_{5\%}$, S_{acin}, S_{cond}, S_{acin}*, S_{cond}*, M1/M0, M2/M0, SIII-DTG z-scores), body plethysmography (FEV_1 and FEF_{25-75} z-scores)	Microbiology (chronic colonisation), treatment (use of antibiotic long-term therapy)	N2-MBW: $LCI_{2.5\%}$ was calculated as the lung volume turnovers required to reach $1/40^{th}$ of the starting N2 concentration. All subjects performed 2 different tidal gas washout measurements, triplicate N2-MBW and DTG-SBWm according to consensus (357). S_{cond} was calculated from the phase III slope (SIII) of washout breaths between the 1.5th and 6th lung turnover. S_{acin} was derived from the first nitrogen SIII and reflects regional acinar ventilation inhomogeneity. LCI5%, S_{cond}(*) and S_{acin}(*) were calculated from abbreviated protocols requiring washout until 1/20th instead of $1/40^{th}$ of the initial nitrogen concentration, and the (*) indices were calculated even earlier. DTG-SBW: SIII was calculated between 65% and 95% of the expired tidal volume and adjusted for tidal volume, as recommended.

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Oktem *et al* (2013) (305)	Body plethysmography (FVC, FEV_1 and FEV_1/FVC % predicted), sleep questionnaire, PSQI (score, poor sleepers, good sleepers), polysomnography, HRCT	Anthropometry (weight and height z-scores)	Severity of symptoms score: cough, sputum production, sputum colour, amount of sputum, wheezing, and breathlessness within the previous month was scores from 0 = none to 3 = severe. Habitual snoring was defined as snoring more than 3 days a week. HRCT: modified Brody score, with the total score derived by adding scores for each abnormality, and ranged from 0 to 37. Pittsburgh Sleep Quality Index (PSQI): "poor sleeper" was defined as those with a score of ≥ 5. Polysomnography: an apnoea hypopnea index of > 1/hr signified a positive polysomnography result and was diagnosed with obstructive sleep apnoea syndrome. Mixed apnoeic events were counted as obstructive. The following parameters were reported: total sleep time in minutes, sleep efficiency (%), Arousal index (n/hr), stage 1 (%TST), stage 2 (%TST), slow wave sleep (%TST), rapid eye movement sleep (%TST), mean saturation (%), mean lowest saturation, obstructive apnoea (n/hr), mixed apnoea (n/hr), hypopnea (n/hr), apnoea–hypopnea index. Sleep questionnaire: habitual snoring, witnessed sleep apnoea, excessive daytime sleepiness, difficulty breathing during sleep, increased parental anxiety about

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			child's sleep, restless sweating, blue colour during sleep, parental shaking for apnoea).

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Olveira *et al* (2017) (210)	Spirometry (FEV_1, FVC, FEV_1/FVC, FEV_1>80%, FEV_1 50%–80% and FEV_1<50% predicted), microbiology (chronic bronchial infection by any pathogen, by *Pseudomonas aeruginosa*, by *Haemophilus influenzae*), treatment (inhaled antibiotics), CT	Anthropometry (BMI)	Microbiology: chronic bronchial infection (CBI) was defined as 3 or more positive cultures for a microorganism in a 6-month period. Spirometry: patients were classified according to their FEV_1 into 3 groups: FEV_1 >80%, between 50%–80% and <50%. Bronchiectasis can be diagnosed from clinical and radiological criteria, bronchography or computed tomography (CT) according to the criteria of Naidich *et al* (358). Bronchiectasis was classified as localized, bilateral, or diffuse (≥ 4 lobes). Patients diagnosed according to clinical-radiological criteria only were excluded.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Paff *et al* (2013) (306)	Exhaled breath profile (volatile organic compounds)	Spirometry (best FEV_1 and FVC in past year % predicted), microbiology (positive bacterial cultures by pathogens), pulmonary exacerbations (number of episodes)	Pulmonary exacerbation: defined as the need to start additional antibiotic treatment as a consequence of a recent change in at least 2 of the following: change in sputum volume or colour, increased cough, increased shortness of breath, increased malaise, fatigue or lethargy, temperature over 38^{o} Celsius, anorexia or weight loss, change in sinus discharge, change in physical findings on examination, decrease in pulmonary function by 10% or more and radiographic changes, according to CBO guidelines based on internationally accepted criteria. Exhaled breath profile: collected with reverse valve system allowing tidal inspiration through a face mask and inspiratory VOC filter and tidal expiration into the spacer. The VOC filter minimizes the influence of environmental VOCs on the breath profile as a potential source of bias. The spacer was connected to the electronic nose during sampling for direct sample analysis during tidal breathing.

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Paff *et al* (2017) (106)	HRQoL (change in SGRQ total score, SGRQ subscores and QoL-B scales), LRTI-VAS (modified score for chest pain), exacerbations (number of pulmonary exacerbations), inflammatory markers in blood (C-reactive protein, erythrocyte sedimentation rate, white blood cell count, neutrophils, eosinophils, basophils, lymphocytes, monocytes),	Anthropometry (BMI), MRC dyspnoea scale score (0-2, ≥3), HRCT or chest radiography (bronchiectasis severity index score: mild, moderate, severe)	HRQoL: change in SGRQ total score (0–100, with 100 being worst QoL) after 12 weeks of treatment was the primary outcome. A 4-point reduction in SGRQ total score has previously been used as the minimal clinically important difference (MCID). Secondary outcomes included sub-scores of the SGRQ and the QoL-B (0–100, with 0 being worst QoL). SGRQ has 50 items with 76 weighted responses divided into 3 categories (symptoms, activity, impact). The categories are scored separately and can be added to provide a total score ranging from 0 to 100, with 0 indicating no impairment of health-related quality of life. The QoL-B is the first disease-specific HRQoL measure for non-CF bronchiectasis patients and includes 37 items on 8 scales (respiratory symptoms, physical, role, emotional and social functioning, vitality, health perception and treatment burden). The scores range from 0-100, with 0 indicating maximum impairment of HRQoL. Minimal clinically important differences range from 7-10 for the different domains. Inflammatory markers: serum C-reactive protein, erythrocyte sedimentation rate, white blood cell count and cell differentiation, microbiological evaluation, sputum cell differentiation, sputum neutrophil elastase, interleukin-1β, -6, -8 and -10, tumour necrosis factor-α, myeloperoxidase, IFN-α and -β. Adherence was determined by the investigator count of all ampoules.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	inflammatory markers in sputum (% sputum cell differentiation, IL-1B, IL-6, IL-8, IL-10, TNF-α, neutrophil elastase, myeloperoxidase, IFN-α, INF-β), spirometry (FEV_1, FVC, FEF_{25-75} % predicted), adverse events, adherence		LRTI-VAS: Symptoms were measured using a modified lower respiratory tract infection visual analogue scale (LRTI-VAS). Four of five symptom domains were scored similar to the LRTI-VAS: dyspnoea, fatigue, cough, chest pain, with sputum colour replaced by ease of sputum expectoration. Pulmonary exacerbation: defined as an acute and significant change in one or more of the common symptoms of bronchiectasis (increase in sputum volume or purulence, worsening dyspnoea, increased cough, declining lung function, increased fatigue/malaise) or the appearance of new symptoms (fever, pleurisy, haemoptysis, requirement for antibiotic treatment), as described by the British Thoracic Society Guideline for non-CF bronchiectasis.
Phillips *et al* (1998) (213)	Spirometry (changes in % in FEV_1 and PEFR in response to exercise and to bronchodilator, baseline measurements FEV_1,	None	Spirometry: baseline pulmonary function was recorded as the best of three flow volume loops. Significant change was 11% for FEV_1, 9% for FVC and 17% for PEFR. Treadmill exercise test: performed according to standardised protocol (359). Bronchodilator response was assessed by giving 200 µg salbutamol via a metered-dose inhaler and spacer device under supervision. PEFR and the best of three flow

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	FVC, FEF_{25-75} and PEFR % predicted)		volume loops were recorded before and 15 min after administration of the bronchodilator.
Pifferi *et al* (2010) (307)	HRQoL (SGRQ and SF-36)	Treatment (daily physiotherapy, regular antibiotics, regular bronchodilators, intermittent bronchodilators, mucolytics, surgical procedures)	HRQoL: SGRQ contains 50 items and 76 weighted responses divided into three components: symptoms, activity and impacts. The symptoms component comprises of eight items concerning the level of symptoms, including frequency of cough, sputum production, wheeze, breathlessness, and the duration and frequency of breathlessness or wheeze. The activity component (16 items) is concerned with physical activities that either cause or are limited by breathlessness. The impacts component (26 items) covers a range of aspects concerning social functioning and psychological disturbances resulting from airways disease. Scores ranging from 0 to 100 are calculated for each component, as well as a total score which summarises the responses to all items. A zero score indicates no impairment of quality of life. The SF-36 questionnaire contains 36 items which provide eight scales, four of which relate to physical health: physical functioning, role physical, bodily pain and general health. The remaining four scales are related to mental health: vitality, social

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			functioning, role emotional and mental health. Each scale is scored from 0 to 100. A score of 100 in physical functioning, role physical, bodily pain, social functioning and role emotional indicates absence of limitations or disability, while in general health, mental health and vitality the best health corresponds to a score of 50. These eight scales provide two summary scores: Physical Component Summary and Mental Component Summary, in which a normal score is 50±10. The normal value is 50 and diminishing scores indicate worsening conditions. A study-specific questionnaire on PCD/Kartagener Syndrome was used comprising of 15 questions relating to diagnosis, clinical features, follow-up, therapy and the presence of other PCD patients within the family. Questions on quality of life improvement after diagnosis were scores from 1=greatly worsened to 5=greatly improved.

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Pifferi *et al* (2012) (228)	Body plethysmography (FEV_1, FVC, FEF_{25-75}, FRCpleth, RV, TLC, RV/TLC, airway resistance (Raw), specific airway resistance (sRaw) and effective specific resistance (sReff) % predicted and z-scores), HRCT	Microbiology (infection with *Pseudomonas aeruginosa*)	Body plethysmography: to be accepted, single inspiratory manoeuvres needed to yield superimposable X-Y plots and values of FRCpleth had to be within 5% of each other. HRCT: Modified Bhalla system, which includes severity of bronchiectasis (score 0-3) and extent of bronchiectasis (score 0-3), mucous plugging (score 0-3), peribronchial thickening (score 0-3), parenchymal abnormalities such as atelectasis (score 0-3) and focal air-trapping (score 0-3). Bronchiectasis was identified according to standard criteria (360). A severity class (from 1 to 3) for total lung impairment was obtained (class of severity 1 for total score of 0-6, class 2 for total score of 7-12, class 3 for total score 13-18).
Pifferi *et al* (2015) (214)	None	Spirometry (FEV_1, FVC and FEF_{25-75} % predicted, changes in FEV_1 and FEF_{25-75} % predicted after bronchodilator), HRCT (bronchiectasis (%), class	Spirometry: best of three flow volume loops was recorded (15 minutes after administration of bronchodilator, when applicable). The % change in FVC. FEV_1 and FEF_{25-75} was calculated to assess bronchodilator response. HRCT: same as above. Secondary ciliary dyskinesia: defined as abnormal ciliary movement or abnormal

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		total lung impairment, class extent of bronchiectasis, class severity of bronchiectasis), microbiology (infection with *P aeruginosa*)	TEM results that are not PCD-specific or that disappear upon cellular regrowth in culture.
Ratjen *et al* (2016) (126)	Inflammatory markers from sputum (IL-8, neutrophil elastase activity, total cell count, % neutrophils, absolute neutrophils, bacterial density), spirometry (FEV_1, FVC and $FEF_{25–75}$ % predicted, change in	None	Pulmonary exacerbation: defined as an increase in respiratory symptoms treated with oral antibiotics. CF Akron pulmonary exacerbation score was used to measure exacerbation severity in patients with PCD and CF (361). Inflammatory markers and microbiology: obtained from spontaneously expectorated sputum.

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	FEV_1 and FVC from baseline in %, pulmonary exacerbation score), microbiology (presence of the pathogens in sputum)		
Rogers *et al* (2013) (223)	Microbiology (bacterial loads, dominant genus relative abundance)	Spirometry (FEV_1 % predicted)	Exacerbations: defined as a change in respiratory symptoms that the PCD specialist considered to be caused by a lower respiratory tract infection requiring antibiotic therapy.
Santamaria *et al* (2008) (227)	HRCT	Spirometry (FEV_1 and FVC % predicted), microbiology	HRCT: Brody score modified to assess the hyperinflation by mosaic perfusion pattern since only the findings of inspiratory CT scans were available for the study (218). Observations were made on six lobes, with the lingula being regarded separately. In patients with situs viscerum inversus, the lung in which the middle lobar bronchus and the corresponding middle lobe was considered as the right lung. A score was calculated for each abnormality, and these scores were summed

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			to provide a total score for each lobe. The scores for the six lobes were then summed to provide a total HRCT scan score, with a theoretical range from 0 (normal) to 243 (maximal score in all lobes). Sub-scores were also calculated for each abnormality by limiting the score to the finding of that abnormality. All scores were normalized to a scale of 0 to 100, representing a percentage of the maximum possible score. A total score of > 5% was abnormal, as in a recent CF study (362). Spirometry: FEV_1 of > 85% predicted was considered normal. Microbiology: deep throat or sputum cultures were obtained.

Appendix C

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Santamaria *et al* (2014) (308)	Respiratory polysomnography (obstructive apnoea index, central apnoea index, hypopnoea index, apnoea–hypopnoea index, oxygen desaturation index, mean oxygen desaturation %, mean and nadir oxygen saturation %), sleep questionnaire (Sleep Disturbances Scale for Children), HRCT	Spirometry (FEV_1, FVC and FEF_{25-75} % predicted), anthropometry (BMI), treatment (n antibiotic courses in the last year), microbiology (positive sputum cultures in the last year)	Respiratory polysomnography: Apnoea–hypopnoea index and oxygen desaturation index (ODI) ≤ 1 per hour were considered normal. Obstructive sleep apnoea syndrome was defined mild, moderate or severe if apnoea–hypopnoea index was >1 to <5, ≥5 to <10, and ≥10, respectively. Sleep questionnaire: Sleep disturbances scale used for school-aged children made of 26 items subdivided into six disorder subscales, i.e. disorders in initiating and maintaining sleep, sleep disordered breathing, disorders of arousal, sleep–wake transition disorders, disorders of excessive somnolence and sleep hyperhidrosis. The total score ranges between 26 and 130, and higher scores indicate more severe disturbances. HCRT: modified Helbich score (220).

Shah *et al* (2016) (309)	Body plethysmography (FEV_1, FEV_1/FVC, TLC, RV/TLC, TLCO and KCO % predicted, estimated change in FEV_1 % predicted per year), HRCT, microbiology (*P aeruginosa* colonisation, non-tuberculosis mycobacteria infection, allergic bronchopulmonary aspergillosis, other pathogens, cumulative sputum analysis)	None	Body plethysmography: lung function at time of diagnosis or transition to adult care was used to determine baseline. Longitudinal lung function data were obtained from patients with at least two lung function records when clinically stable with a minimum of three forced expiratory manoeuvres within the same lung function laboratory in the absence of bronchodilator. Lung function decline was expressed as FEV_1 % predicted and estimated using Global Lung Function Initiative reference equations. Microbiology: chronic colonisation was defined as the isolation of potentially pathogenic bacteria or fungi in the sputum on two or more occasions at least 3 months apart in a 1-year period with >50% positive cultures during the year. All patients had three or more sputum cultures over the duration of follow-up. Sputum microbiology for patients was presented as cumulative colonisation over the duration of the follow-up period. Nontuberculous mycobacteria infection was defined according to the ATS guidelines and allergic bronchopulmonary aspergillosis according to the British Thoracic Society guidelines. Body plethysmography: European Community for Steel and Coal reference equations were used for measurement of transfer factor of the lung for carbon monoxide (TLCO). HRCT: extent of bronchiectasis, severity of bronchial dilatation, bronchial wall thickness, mucus plugging in large and small airways, mosaicism and emphysema were scored for each lung lobe (the lingula was considered as a different lobe,

			making a total of 6 lobes), according to a modified Bhalla system (219). The scoring system was as follows: 1) extent of bronchiectasis (0 = none, 1 = one or partial bronchopulmonary segment involved, 2 = two or more bronchopulmonary segments involved, 3 = generalized cystic bronchiectasis); 2) severity of bronchial dilatation (0 = normal, 1 = less than twice the diameter of the adjacent pulmonary artery, 2 = more than twice the diameter of adjacent pulmonary artery); 3) severity of bronchial wall thickening (0 = normal, 1 = <0.5 × the diameter of the adjacent pulmonary artery, 2 = 0.5 - 1.0 × the diameter of the adjacent pulmonary artery, 3 = ≥ 1.0 × the diameter of the adjacent pulmonary artery); 4) presence of mucous plugging in large airways (0 = none, 1 = minimal, 2 = extensive 5) presence of mucous plugging in small airways (0 = none, 1 = minimal, 2 = extensive); 6) extent of mosaicism (to nearest 5%) and 7) extent of emphysema (to nearest 5%). Patients with previous lobectomies had scores adjusted to represent the maximum score available. Scores for extent of bronchiectasis, severity of bronchial dilatation and thickening and mucus plugging in small and large airways are expressed as percentages of maximum possible score.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Shoemark *et al* (2009) (212)	FENO (FENO50, FENO100, FENO200, J'awNO, CalvNO)	Anthropometry (height, weight), spirometry (FEV_1 raw), treatment (requirement for antibiotics, inhaled corticosteroids), microbiology (pathogens in sputum samples), nasal NO (ppb)	Fraction of exhaled nitric oxide (FENO): J'awNO is total NO flux in the airways and CalvNO is steady-state NO concentration in alveolar air. The mean of 2 FENO measurements at each flow rate measured (50, 100 and 200 ml/s) was used to calculate J'awNO and CalvNO, according to ATS standards. Nasal NO: measured according to ATS/ERS standards using the breath-hold technique for velum closure (363).
Smit *et al* (1996) (310)	Lung resection (location and extent), symptoms questionnaire	Spirometry (FEV_1 and FVC % predicted), bronchiectasis (n and % bilateral), dyspnoea index (0+1, 2, 3+4), hospitalisations	Symptoms questionnaire: present complain about daily cough, phlegm, haemoptysis, respiratory infections, dyspnoea, fitness for work, and the influence of resection on pulmonary complaints.

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Sunther *et al* (2016) (125)	Spirometry (FEV_1 % predicted, baseline FEV_1 < 40%, mean baseline and admission FEV_1 % predicted)	Anthropometry (BMI), microbiology (persistent infection with pathogens), treatment (*n* treated with intravenous antibiotics, *n* oral prophylactic antibiotics, *n* in use of hypertonic saline or rhDNase)	Pulmonary exacerbation: defined as change in respiratory status for which intravenous antibiotics were prescribed. Spirometry: FEV_1 % predicted values were calculated using the Global Lung Initiatives (GLI) equations. Baseline FEV_1 was defined as the best FEV_1 in the 12 months before the pulmonary exacerbation. Recovery to baseline was defined as any FEV_1 within 3 months after treatment that was greater than or equal to 90% of the baseline FEV_1. Microbiology: persistent infection was defined as at least two positive growths of the same microorganisms on cough swab or sputum culture in the 12 months before the pulmonary exacerbation.
Tamalet *et al* (2001) (312)	Spirometry (FEV_1 % predicted), blood gas (mean arterial PO_2)	CT (bronchiectasis, radiologic deterioration, lobectomy), treatment (antibiotic use)	Respiratory tract infections: defined as persistent cough with bronchial rhonchi, with or without fever. Frequency of infections: classified as less than or more than 6 infections per year since birth. CT: presence of bronchiectasis (internal diameter of bronchus larger than that of an adjacent artery) was assessed, and its topography was scored as absent, unilateral, or bilateral. The course of bronchiectasis was evaluated by CT scan performed

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			every 2 years and classified as stable or progressive. Radiologic deterioration corresponded to bronchiectasis extension. Blood gas: arterialized capillary blood. Spirometry: results were expressed as a percentage of the expected value for age and considered as normal when > 80% of the expected value. Pulmonary function tests were performed at least twice in 35 of 41 children, at a mean interval of 6 years. Treatment: frequency of antibiotic use prescribed over the entire follow-up period for their lower or upper respiratory tract infections was evaluated and scored (no antibiotics, intermittent or continuous).
Valerio *et al* (2012) (313)	Spirometry (FEV_1, FVC and FEV_1/FVC % predicted), exercise test (VO_{2peak}, VE/VCO_2 slope, O_2 pulse, heart rate peak), physical activity assessment	Anthropometry (BMI and BMI SDS)	Spirometry: according to standard spirometric techniques (326). FEV_1 > 85% predicted was considered normal. Physical activity assessment: modified version of the long International Physical Activity Questionnaire for adolescents. The questionnaire focuses on 4 domains: school-related physical activity, including activity during physical education classes and breaks, transportation, housework and leisure time. For each of the 4 domains, the number of days per week and the number of physical activity periods per day (>

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	(total time spent in physical activity, vigorous physical activity)		10 min of walking, moderate activity or vigorous activity) were recorded. Outcome measures were average minutes per day of walking, moderate or vigorous activities, with the sum of these variables computed to obtain minutes per day of total physical activity. Cardiopulmonary exercise test: peak oxygen consumption (VO_2peak) was recorded as the mean value of VO_2 during the last 20 seconds of the test and was expressed in millilitres per kilogram per minute. VO_{2peak} was compared with maximal predicted VO_2 by use of a sex-, age-, height- and weight-adjusted and protocol-specific formula.
Vallet *et al* (2013) (314)	Spirometry (FEV_1, FVC and FRC % predicted, *n* abnormal FRC and FEV_1), blood gas (PaO_2, *n* hypoxemic patients), CT (bronchiectasis, progressive bronchiectasis)	None	Spirometry: at least 3 curves reproducible for FEV_1 were recorded and the best curve was retained for analysis. Flows were considered normal when > 80% of the expected value. Blood gas: arterialised capillary blood gases for hypoxemia, which was defined as a value of PaO2 below the lower limit of normality (2 standard deviations below predicted measures in age-matched healthy children). CT: bronchiectasis was classified as stable or progressive (increasing diameter

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			and/or extension to a new segment). Radiological deterioration was defined as the extension of bronchiectasis.

Appendix C

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Wells *et al* (2011) (315)	Exercise testing (maximal aerobic capacity, maximal oxygen uptake, change in pH after exercise, Pi/PCr ratio (ADP ratio), halftime of PCr recovery in seconds, work during exercise trial in Watts)	Spirometry (FEV_1 and FVC % predicted), anthropometry (height, mass, lean body mass), Habitual Activity Estimation Scale questionnaire	Spirometry: according to standard spirometric techniques (326) and expressed as % predicted value for height and gender (364). Habitual Activity Estimation Scale questionnaire: was used as an estimation of activity levels as previously described and validated in this population (365, 366). Change in pH after exercise (rest pH - end-exercise pH) - Intracellular pH was calculated for each spectrum based on the chemical shift difference between PCr and Pi. The cytosolic [Mg2] was calculated from the chemical shift of ATP measured from the resonance of PCr, and this information was used to correct calculated pH for changes in [Mg2] Halftime of PCr recovery (seconds): The time constant of the recovery rate of PCr was calculated during recovery after each exercise bout using an exponential curve fit Work during exercise trial (Watts): Watts and repetitions per minute (rpm) of the ergometer were recorded every 5 seconds during exercise

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Whalley *et al* (2006) (316)	HRQoL (Stigma score, SGRQ scores on symptoms, activity and impact, SF-36 component scores on physical and mental, questionnaire on mental and physical health status)	None	Stigma rating: each participant was rated on a four-point scale for perceived stigma (1 = no perceived stigma to 4 = high perceived stigma). These rating were based upon an informal subject analysis of psycho-social themes within the qualitative data, including self-reported symptom concealment, trust in medicine, and current and past social support.
Yiallouros *et al* (2015) (317)	None	CT (presence of bronchiectasis), microbiology (presence of pathogens in sputum culture), spirometry (FEV_1 and FVC z-scores, % with low FEV_1 and % with low FVC), anthropometry (BMI	Spirometry: z-scores < -1.96 were considered abnormal. Anthropometry: BMI was expressed as age- and gender-specific z-scores based on the US Centers for Disease Control 2000 growth charts.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		z-score), lobectomy (location of resected lobe)	
Zihlif *et al* (2005) (318)	Cough frequency (n cough episodes), cough symptom score	Spirometry (FEV_1 % predicted), eNO, inflammatory markers (sputum neutrophil count), microbiology (presence of pathogens)	Exhaled Nitric Oxide (eNO): the mean value out of three correctly executed exhalations was recorded. Spirometry: at least 2 manoeuvres were required to have an FEV_1 within 10% of each other. Baseline FEV_1 was recorded as the best of three manoeuvres. Values were expressed as percent of predicted normal values (332). Cough frequency: cough was identified by 2 signals: the electromyography signals from the muscles of active expiration, and a filtered audio signal. Visual inspection confirmed that all cough epochs identified automatically were in fact genuine. Coughing events were counted both as individual spokes and as clusters. Each cluster (cough epoch) was arbitrarily defined as a close succession of cough spikes (<2 seconds between individual coughs) recorded by each trigger of the recorder. Cough data were expressed as total numbers of cough episodes (individual spikes + cough cluster) per recording time.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			Cough symptom score: questionnaires handed to parents, with scores ranging from 0 = no cough to 5 = distressing cough
Zihlif *et al* (2006) (319)	Inflammatory markers from exhaled breath condensate and sputum (IL-8, LTB4 and 8-isoprostane, sputum neuthrophil count)	Spirometry (FEV_1 % predicted)	Stable pulmonary disease: defined clinically as no hospitalisation or changes in antibiotic regimen within 2 weeks prior to being in the study and FEV_1 within 10% of best recorded value in the last year. The volume loop with the highest FEV_1 was selected as opposed to the more conventional sum of FEV_1 and FVC as PCD patients often terminated their expiratory effort by coughing before their residual volume was reached. Sputum: neutrophil cell count was expressed as a percentage of total cell count.

C.3 Computed tomography scoring scales

Table 23 Comparison between computed tomography (CT) scoring scales used in PCD studies included in the clinical outcome measure scoping review.

	Bhalla *et al* (219)	**Brody *et al*** (218)	**Helbich *et al* (original or modified)** (220)
Target population	Developed for cystic fibrosis	Developed for cystic fibrosis	Developed for cystic fibrosis, modified from Bhalla score
CT technique	Interrupted or volumetric	Interrupted or volumetric; requires expiratory images for air-trapping	
Features evaluated	1. Bronchiectasis	1. Bronchiectasis	1. Bronchiectasis
	2. Peribronchial thickening	2. Mucous plugging	2. Peribronchial wall thickening
	3. Mucous plugging	3. Peribronchial thickening	3. Mucous plugging
	4. Sacculations or abscesses	4. Parenchyma (dense opacity, ground-glass opacity, cysts or bullae)	4. Sacculations or abscesses
	5. Bullae	5. Hyperinflation	5. Bullae
	6. Air trapping/ emphysema/ mosaic perfusion		6. Emphsema

	Bhalla *et al* (219)	**Brody *et al*** (218)	**Helbich *et al* (original or modified)** (220)
	7. Collapse/consolidation		7. Collapse/Consolidation
			8. (Mosaic perfusion)
Assessment of severity	Mild / Moderate / Severe for bronchiectasis and peribronchial thickening	Severity multipliers for bronchiectasis size (0-3), severity of bronchial wall thickening (0-2) and type of air-trapping	Severity score (0-3; absent/ mild / moderate / severe) for bronchiectasis, bronchial wall thickening, bullae, emphysema, collapse/ consolidation
Assessment of extent	By number of bronchopulmonary segments involved, and by generations of bronchial divisions	By proportion of lobe involved. Separate assessments of peripheral and central lung.	By number of bronchopulmonary segments involved and by generation of bronchial divisions
Maximum score	Seven features can be scored 0-3, two features can be scored 0-2	Theoretical: 40.5/lobe; 243	Seven features can be scored 0-3 and two (three for modified version) features can be scored 0-2
	Maximum score: 25	Maximum score: 207	Maximum score 25 (27 for modified version)
	Subtract score for each feature from 25 to obtain final patient score	Normalized to 0-100 scale	Subtract score for each feature from 25 (27 for modified) to obtain final patient score

Appendix D Comparison of service delivery models for PCD: study pack documents

D.1 Participant information sheet

Delivery of care for PCD across Europe

Participant Information Sheet

Researcher: Bruna Rubbo Supervisor: Prof Jane Lucas (ERGO number: 20177)

We would like to invite you to take part in a research study looking at the differences in delivery of care for primary ciliary dyskinesia (PCD) across Europe. Please read this information carefully before deciding whether to take part in this research. If you are happy to participate you will be asked to sign a consent form or give verbal agreement to each item on the consent form prior to taking part on this study.

What is the research about?

The aim of this study is to investigate variations in models of service delivery for patients with suspected or confirmed diagnosis of PCD within and between European countries. Knowing how patients are referred, diagnosed and managed through Europe will inform us of intrinsic differences that can and should be taken into account when designing research studies based on international collaboration. This information is also essential to facilitate data collection and sharing across Europe and to ensure these are done in a standardised manner. Additionally, data derived from the study will highlight differences in data availability throughout European countries.

This study is a part of the work of a PhD student (Bruna Rubbo) and will contribute to 'Workgroup 3. Clinical Care' of the BEAT-PCD Project (COST-Action grant BM1407). The study is supervised by Prof Jane Lucas.

Appendix D

Why have I been chosen?

You've been selected to take part because, as a PCD specialist, we believe you can provide essential information for the development of the survey. Your unique knowledge in delivery of care for PCD patients will be important for us to gain a deeper understanding of the different processes and structures of service delivery that are particular to your country.

What will happen if I take part?

This study is comprised of two parts. The first part, which we are inviting you to take part at this time, consists of a short questionnaire and interviews with key healthcare specialists from different European countries. These interviews will provide information that will be used to develop an international survey, taking into account contextual evidence (*e.g*. referral process, diagnostic methods, treatments) that may vary between countries. The survey will be the second part of this project and will be circulated to all healthcare staff involved in care for patients with suspected PCD via the BEAT-PCD network.

If you agree to participate in the study, you will be asked you to then either: a) print this form, initial each box, sign at the bottom and send it to the researcher (scanned copy by email or original copy through post); OR b) email the research team stating that you have read the consent form and agree to participate in the study, in which case you will be asked to verbally confirm your agreement to each item at the start of your interview .You can reach us if you have any questions or would like to discuss any item of the consent form using the contact details provided at the bottom of this information sheet.

After receiving your consent, we will send you a short questionnaire by email, which should take approximately 10 minutes to complete. Once we receive your completed questionnaire, we will contact you to arrange a convenient time to conduct the interview. This can be done face-to-face or through Skype. The interview will last approximately 1 hour, and you will be asked to answer a few questions, which mainly relate to your experience in delivering care for patients with suspected or confirmed PCD. You will have an opportunity to ask your interviewer questions at the start or end of the interview process.

Are there any benefits in my taking part?

There are no direct benefits for you if you decide to part in this study. However, we believe that gaining knowledge on delivery of care models across Europe will facilitate and enhance research in PCD, which you and your research group might already be involved.

Are there any risks involved?

There are no foreseeable risks involved in taking part in this research.

What happens if something goes wrong?

If you have a complaint, please contact the University of Southampton Research Governance Office on (+44) 23 8059 5058 or rgoinfo@soton.ac.uk.

Will my participation be confidential?

All information collected during the interview will be kept strictly confidential. We will ask that you allow us to use a recording device to record your interview so that we do not miss any of the information you give us, and recording will be deleted after transcription. If we decide to use any quote you provide during the interview, this will be done in an anonymised way. You will also have the opportunity to review your interview transcript, if you wish to.

What happens if I change my mind?

You are free to withdrawal at any time, without providing any information. You can contact the study team and request to be taken off the study and to have your data destroyed by using the information provided at the bottom of this information sheet.

How can I contact the study team?

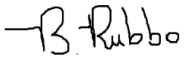
If you have any questions or concerns about this study, please contact the study team using the information provided below.

Bruna Rubbo (b.rubbo@soton.ac.uk; Skype ID: bruna.rubbo); LF100, South Academic Block, Southampton General Hospital, Tremona Road; Southampton SO16 6YD

Prof. Jane Lucas (jlucas1@soton.ac.uk)

Thank you for taking the time to read this information sheet. I hope you will agree to help us by taking part in this project.

Yours sincerely,



BEAT-PCD
BETTER EXPERIMENTAL APPROACHES TO TREAT PCD

D.2 Participant consent form

Researcher: Bruna Rubbo Supervisor: Prof. Jane Lucas

Study ID (to be completed by researcher):

Please read this form carefully. If you agree to participate, please place your initials in each of the boxes

1. I confirm that I have read and understood the Participant Information Sheet entitled 'PIS_v1.0_060616' dated 06/06/2016 and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw from the study at any time without giving any reason by contacting the research. I understand that if I decide to withdraw, all data from my interview will be destroyed, including any audio recordings. ☐
3. I understand that my data will be anonymised and therefore any comments I provide to the interviewer will not be identifiable to anyone outside the immediate research team. My data will be kept in compliance with the Data Protection Act 1998. ☐
4. I agree to complete the short questionnaire entitled 'Questionnaire_v1.1_080716'. ☐
5. I agree to having my interview recorded using an audio recording device (*e.g.* tape recorder). I understand I can request to review the interview transcript by contacting the research team. ☐
6. I agree that the information I provide on the interviews will be used exclusively for research. ☐
7. I agree to take part in this study. ☐

_____	_____	_____
Participant's full name	Date	Participant's signature
_____	_____	_____
Researcher's full name	Date	Researcher's signature

Appendix E Comparison of service delivery models for PCD: interview guide and complete survey

E.1 Interview guide

Thank you very much for agreeing to participate in this study and for completing the short questionnaire. I would like to start by stating the purpose of this study. As I mention in the Participant Information Sheet, this will be a part of my PhD project, which will look at variations in service delivery for PCD across different countries. The aim of this interview is to gain a deeper understanding of service delivery in your country to inform the development of a survey on delivery of care for PCD. It is a project within BEAT-PCD (COST Action BM1407).

It is important that we design the survey in a way that is relevant to different settings. In order to do that we need to gather as much information on how the system works in your country. We are fully aware that it might differ dramatically from the UK setting, for example and we want the survey questions and categories to reflect these differences, which are very important for a better understanding of service delivery methods for PCD in different countries.

The interview will take roughly an hour so I would like to confirm with you that we do have sufficient time for this. If something urgent comes up during the interview, please let me know and we can arrange a convenient time to resume the interview.

Do you have any questions for me before we start?

Thank you for completing and returning the questionnaire. It provides me the basic information about you and your work.

1) Could you please describe your role in PCD diagnostics and or management?
2) Could you describe the service provided for PCD patients in your country?
3) How do patients with suspected PCD reach you?
4) What would you say are the main barriers for PCD patients to reach your service?
5) What do you think would be an appropriate definition for a PCD specialist centre?
6) Applying that definition, would you say you work in one?
7) Could you describe your workplace?
8) What do you think or expect will change if the European Reference Networks come into play?
9) What impact do you expect these networks will have on your service?

10) (if yes to **visiting other PCD specialist** centre) You mentioned in the questionnaire that you have visited other PCD specialist centres. Did it differ in any way to your workplace? What were the differences you observed?
11) Could you describe what happens when patients come in for an appointment?
12) You said you regularly use tests A, B, C, etc for diagnostic purposes. Could you please tell me more about them?
13) How do patients get diagnosed with PCD? How is a diagnosis reached?
14) Is there a different process for cases that are difficult to diagnose?
15) What happens if a definite diagnosis cannot be reached?
16) Could you describe what happens when patients come in for an appointment?
17) You said you regularly use tests (A, B, C, etc from questionnaire) for diagnostic purposes. Could you please tell me more about them?
18) How do patients get diagnosed with PCD? How is a diagnosis reached?
19) Is there a different process for cases that are difficult to diagnose?
20) What happens if a definite diagnosis cannot be reached?
21) How adequate would you say is this transition process?
22) Could you explain how the PCD diagnostics and management (or your) services are funded?
23) Could you please tell me a bit more about funding for treatments such as medications and physiotherapy?
24) You mentioned on the questionnaire that you or your service is involved in PCD-related research. Could you tell me more about this?
25) What do you think are the most important aspects in delivery of care in PCD?
26) How important do you think are PCD specialist centres?
27) What do you feel has changed since the implementation of specialist centres?
28) What do you think of PCD patient organisations?
29) What is your opinion on the use of telemedicine such as Skype, telephone calls and video conferencing?
30) What is your opinion on the establishment of disease registries?
31) How can we improve data collection and quality?
32) How can we improve delivery of care for PCD in your country?
33) How can we improve delivery of care for PCD internationally?

Before we conclude the interview, would you like to comment on any other points that we did not cover during this interview? Is there anything you would like to ask me?

Once again, I want to thank you for your participation in this project. Your views and opinions are extremely important, and I appreciate you taking the time to help us better understand the differences between delivery of care for PCD in different countries.

E.2 Service delivery complete survey

Please ensure you have read the Participant Information Sheet (dated 12/03/2019 v1.2) carefully and signed the consent form (dated 12/03/2019 v1.2) before completing this survey. If you have any questions about this study, please contact the research team using the information provided in the Participant Information Sheet.

Please read the following instructions carefully before completing the survey:

- **All questions refer to your service and your opinions on your service.** If your institution has a separate adult and paediatric service, please answer the questions according to the service you are primarily involved.
- Answer all questions to the best of your ability and **provide estimates/your best guess** when you are not sure of the exact numbers.

Thank you for taking the time to complete this survey.

DEMOGRAPHICS

Country:

_____

City:

_____

Institution:

_____

_____

_____

Position/title:

_____

_____

Speciality:

_____

Date of survey completion:

_____

Age: ≤25 years ☐ 26-34 years ☐ 35-44 years ☐ 45-54 years ☐ ≥55 years ☐

Appendix E

Sex: Male ☐ Female ☐

Which career stage do you identify with? Early stage ☐ (e.g. PhD student, pre-registrar) Mid-stage ☐ (e.g. Post-doc, registrar) Advanced stage ☐ (e.g. Professor, group lead, consultant)

Predominant duties within the PCD team, as defined by your job role. **Please select only one option.**

Diagnostics	☐
Management	☐
Laboratory researcher	☐
Clinical researcher	☐

If you have another predominant duty, please specify_____

Other duties within the PCD team. **Please tick all that apply.**

Diagnostics	☐
Management	☐
Laboratory researcher	☐
Clinical researcher	☐

If you have another duty, please specify_____

Which of the following clinical activities are you involved with? **Please select all that apply.**

Children with PCD	☐
Adults with PCD	☐
No clinical duties	☐

Which of the following research areas are you actively involved in? **Please select all that apply.**

PCD	☐
CF	☐
Non-CF bronchiectasis	☐
ILD	☐
Asthma	☐
Other rare disease	☐
Other	☐

If other, please specify _____

PART A. COMPARISON WITH 2007-2009 EUROPEAN SURVEY

1. Type of institution you work in. **Please select all that apply.**

University hospital	☐
General hospital	☐
Private hospital	☐
Private practice	☐
Other	☐

2. Characteristics of your institution. **Please select only one option.**

Primary care centre (paediatric practice or small hospitals)	☐
Secondary care centre (regional referral centre with respiratory unit)	☐
Tertiary care centre (university hospital or other tertiary referral centre)	☐

3. **Number of PCD patients diagnosed and treated in this centre.**

On average, how many cases are referred for diagnostic evaluation every year? ☐

On average, how many new PCD cases are diagnosed every year? ☐

Total number of PCD patients who are being cared for at the centre (and were seen at least once between Jan 1, 2017 and Dec 31, 2018?) ☐

4. If you see a new patient with suspected PCD, which diagnostic procedures do you perform? **Tick all that apply.**

Test	**Performed in centre**	**Not performed in this centre but sent to:**		**Not performed**
		National centre	**International centre**	
Nasal nitric oxide	☐	☐	☐	☐
Nasal biopsy/brushing	☐	☐	☐	☐
Bronchial biopsy	☐	☐	☐	☐
Ciliary beat frequency	☐	☐	☐	☐
Ciliary beat pattern (using high-speed video microscopy)	☐	☐	☐	☐
Electron microscopy	☐	☐	☐	☐
Cell culture	☐	☐	☐	☐
Genetic testing	☐	☐	☐	☐
Immunofluorescence	☐	☐	☐	☐
Saccharine test	☐	☐	☐	☐
Other	☐	☐	☐	☐

5. Once you have established a diagnosis of PCD, how often do you usually see the patients for follow-up visits?

Frequency	**seen by respiratory physician**	**seen by physiotherapist**	**seen by ENT specialist**
Every 3-4 months	☐	☐	☐
Every 6 months	☐	☐	☐
Every year	☐	☐	☐
Every 2 years	☐	☐	☐
At patient/parent's request	☐	☐	☐
Do not know	☐	☐	☐

6. Which therapies do you usually or sometimes prescribe for your patients with PCD?

Therapy	**Routinely**	**Frequently**	**Rarely**	**Never**
Airway clearance	☐	☐	☐	☐
Formal exercise programme	☐	☐	☐	☐
Antibiotics for exacerbations	☐	☐	☐	☐
Antibiotics if pathogens found on culture	☐	☐	☐	☐
Prophylactic antibiotics	☐	☐	☐	☐
Intermittent intravenous antibiotics	☐	☐	☐	☐
Regular intravenous antibiotics	☐	☐	☐	☐
Inhaled corticosteroids	☐	☐	☐	☐
Bronchodilators	☐	☐	☐	☐
rhDNase	☐	☐	☐	☐
Hypertonic saline	☐	☐	☐	☐
Tympanostomy/tube insertion	☐	☐	☐	☐
Nasal rinsing	☐	☐	☐	☐
Vaccination (i.e. Influenza)	☐	☐	☐	☐
Other	☐	☐	☐	☐

If other, please specify _____

PART B. YOUR SERVICE

7. Members of your PCD Team. Please complete each box with the **number of professionals** available in your service or select not available, if professional not available at your service

Roles	Core PCD team	PCD team has access to	Not available
Respiratory paediatrician	☐	☐	☐
Adult pulmonologist	☐	☐	☐
ENT surgeon	☐	☐	☐
Audiologist	☐	☐	☐
Specialist nurse (e.g. PCD, CF)	☐	☐	☐
Non-specialist nurse	☐	☐	☐
Specialist physiotherapist	☐	☐	☐
Non-specialist physiotherapist	☐	☐	☐
Physiologist	☐	☐	☐
Social worker	☐	☐	☐
Psychologist	☐	☐	☐
Nutritionist	☐	☐	☐
Geneticist	☐	☐	☐
Fertility specialist	☐	☐	☐
Other	☐	☐	☐

If other, please specify _____

8. Which of the below describe the role of the PCD multidisciplinary team (MDT) in your centre? **Please tick all that apply.**

No MDT available at the centre	☐
All diagnostic decisions are reached with input from MDT	☐
Only difficult diagnostic cases are discussed with MDT	☐
All patients are managed with input from MDT	☐
Only difficult management cases are discussed with MDT	☐

Please expand on the role of your MDT: _____

_____

9. Is there a patient support organisation in your country?

No	☐
Yes	☐
Do not know	☐

Referrals

10. How often are patients referred to the centre for diagnostic testing from each of the following sources?

Source	**Mostly**	**Occasionally**	**Never**
Internal referrals (e.g. from another department)	☐	☐	☐
General practitioners	☐	☐	☐
Paediatricians	☐	☐	☐
Paediatric pulmonologists	☐	☐	☐
Adult pulmonologists	☐	☐	☐
Ear, nose and throat specialist	☐	☐	☐
Neonatologists	☐	☐	☐
Cardiologists	☐	☐	☐
Bronchiectasis clinics	☐	☐	☐
Self-referred (i.e. no referral from clinician)	☐	☐	☐
Fertility clinic	☐	☐	☐
Other	☐	☐	☐

Please specify other source, if applicable _____

11. Does your centre receive referrals from patients from other centres? **Please tick all that apply.**

No	☐
Yes, we receive referrals of patients **to be diagnosed** at our centre	☐
Yes, we receive referrals of patients **diagnosed** at other centres in **our country**	☐
Yes, we receive referrals of patients **followed-up** at other centres in **our country**	☐
Yes, we receive referrals of patients **diagnosed** in **another country**	☐
Yes, we receive referrals of patients **followed-up** in **another country**	☐

PCD diagnostic centre

11. Total number of PCD diagnostic centres available in your country (including your centre)

☐

12. How would you describe your PCD diagnostic service?

Centralised	☐
Decentralised	☐

PCD management centre

13. Where are patients **usually** followed-up?

Clinics exclusively for PCD patients	☐
CF clinics	☐
Bronchiectasis clinics	☐
Other	☐

If other, please specify _____

14. How are patients **usually** followed-up?

Exclusively at PCD centre	☐
Shared care with local clinicians	☐

15. Total number of PCD management centres available in your country (including your centre):

☐

16. How would you describe your PCD management service?

Centralised	☐
Decentralised	☐

Funding

17. Please select **the option that best describes** the type of healthcare system in your country.

Public (government funding)	☐
Private (healthcare insurance)	☐
Mostly public with some private	☐
Mostly private with some public	☐
Half public and half private	☐

18. Please select the **option that best describes** the healthcare coverage available to the **majority of patients** that reach your centre.

Public healthcare coverage	☐
Private healthcare insurance	☐

19. Which of the following activities is **covered by government funding in your country** (e.g. through taxes or government or compulsory health insurance)?

Type of activity	**Full coverage**	**Partial coverage**	**No coverage**
Outpatient appointment	☐	☐	☐
Hospitalisations	☐	☐	☐
Diagnostic testing (excl. genetics)	☐	☐	☐
Genetic testing	☐	☐	☐
Low-cost medications (e.g. antibiotics)	☐	☐	☐
High-cost medications (i.e. rhDNase)	☐	☐	☐
Physiotherapy equipment in-hospital	☐	☐	☐
Elective surgery	☐	☐	☐
Emergency visits	☐	☐	☐
Regular patient transport to/from centre	☐	☐	☐

20. Are there any additional provisional funding from **the government** if patients are diagnosed? **Please select the option that best describes your service.**

All or a portion of the treatment is free only if patient is **diagnosed with PCD**	☐
All or a portion of the treatment is free only if patient is **diagnosed with a rare disease**	☐
Treatment is **free**, independent of diagnosis	☐
Treatment is **not free**	☐

21. Where does funding for the PCD service come from? **Please tick all that apply.**

International funding (e.g. European)	☐
National funding	☐
Local/regional funding	☐
Research funds	☐
Do not know	☐
Other	☐
If other, please specify	

22. Please describe, in your own words, the funding structure for PCD services in your country.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

PART C. OPINIONS ON DELIVERY OF CARE FOR PATIENTS WITH PCD

23. What are the **main barriers** for patients to reach the PCD service for **diagnostic assessment**? **Please rank** each of the following items in order of importance, with 1 being the most important to 5 being the least important.

Rank from **1 (most important) to 5 (least important). Assign each number to one box only.**

Patients travelling long distances to reach the centre	☐
Lack of awareness of PCD by clinicians	☐
Lack of awareness of PCD by people with symptoms or family members	☐
No clear referral pathway (i.e. where to refer patients for testing)	☐
Patients not seeking help for their symptoms	☐

24. What can we do to **improve access to service for diagnostics**? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 7 (least important). Assign each number to one box only.**

Engage with patient organisations	☐
Raise awareness of disease among clinicians	☐
Raise awareness among patients of the need to seek diagnostic services	☐
Wider dissemination of guideline on who needs to be referred for testing	☐
Promote international collaboration for cross-border diagnostic testing	☐
Move towards a centralised system with limited PCD centres concentrating resources	☐
Move to a decentralised system for better access to PCD services	☐

25. What are the **current challenges for PCD diagnostics**? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 7 (less important). Assign each number to one box only.**

Lack of standardisation for diagnostic testing and reporting	☐
Difficulties in adopting and/or adhering to published diagnostic guidelines in your own setting	☐
Delay between date of testing and availability of results	☐
Inconclusive and difficult to diagnose cases	☐
High cost of diagnostic equipment	☐
Need to send samples to another country/centre for diagnostic testing	☐
Lack of appropriate training to conduct and interpret diagnostic test results	☐

26. Which option **best describes** your views on the **ideal number of PCD diagnostic centres per country**? **Please select only 1 option.**

Number of expected PCD cases (based on national population and disease incidence)	☐
One centre per region (based on geographical distribution)	☐

27. What are the **main barriers** for patients to reach the PCD service for **follow-up clinics**? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 5 (less important). Assign each number to one box only.**

Patients travelling long distances to reach the centre	☐
Lack of understanding by clinicians of severity of the disease and need for regular follow-up	☐
Lack of understanding by patients of severity of the disease and need for regular follow-up	☐
No evidence-based PCD therapies available	☐
Limited funding	☐

28. What can be done to **improve follow-up of PCD patients**? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 7 (less important). Assign each number to one box only.**

Raise awareness of severity of disease among clinicians	☐
Raise awareness of severity of disease among patients	☐
Promote international collaboration for cross-border care	☐
Move towards a centralised system with limited number of PCD centres	☐
Move towards a decentralised system with small local PCD centres	☐
Favour a shared care system between PCD specialists and local clinicians	☐
Implement telemedicine for following-up patients that live far from PCD centre	☐

29. What are the **current challenges for PCD management in general**? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 5 (less important). Assign each number to one box only.**

Patient education on severity of PCD	☐
Loss-to-follow-up due to decentralised care	☐
Lack of evidence-based guidelines for PCD management	☐
Patients with mild disease do not want to travel long distances for regular follow-ups	☐
Centres with few patients not having sufficient expertise to treat PCD	☐

30. What are the **current challenges for adult PCD management**? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 4 (less important). Assign each number to one box only.**

Loss-to-follow-up due to absence of adult PCD service	☐
Lack of resources (e.g. personnel, rooms, funding)	☐
Adult clinicians often care for patients with more severe disease so PCD is not their priority	☐
Lack of formal transition from paediatric to adult care	☐

31. Which option **best describes** your views on the **ideal number of PCD management centres per country**? **Please select only 1 option.**

Number of expected PCD cases (based on disease prevalence in the region)	☐
One centre per region (based on geographical distribution)	☐
Management centres should always be combined with a PCD diagnostic centre, so number of management centres depends on number of diagnostic centres	☐

32. When setting up a **new PCD centre**, how would you prioritise the following items? **Please rank** each of the following items in order of importance.

Rank from **1 (most important) to 5 (less important). Assign each number to one box only.**

Awareness of any/all existing PCD centres in the country	☐
Collaboration with well-established centres to send samples for diagnostic testing	☐
Obtain support from national or regional government bodies	☐
Travel to well-established centres to observe their setting	☐
Being a member of PCD networks	☐

33. When setting up a **new PCD centre**, how would you prioritise the following **local resources**? Please rank each of the following items in order of importance.

Rank from **1 (most important) to 5 (less important). Assign each number to one box only.**

Awareness of latest research findings	☐
Availability of multidisciplinary team	☐
Expert personnel, trained at well-established centres	☐
Availability of diagnostic tests to establish a diagnosis of probable PCD (nNO, HSVA)	☐
Good relationship with local colleagues	☐

34. When defining a **PCD specialist centre,** how important are each of the following items. **Please rate** each of the items from 1 (not important at all) to 5 (very important).

1: Not at all important; 2: Slightly important; 3: Neutral; 4: Important; 5: Very important

Having diagnostic and management combined in one centre	1	2	3	4	5
Expertise in conducting and interpreting diagnostic tests	1	2	3	4	5
Capability to conduct all diagnostic tests	1	2	3	4	5
Capability to conduct diagnostic tests that can confirm PCD diagnosis (i.e. genetics, TEM)	1	2	3	4	5
Being engaged in PCD research	1	2	3	4	5
High throughput of patients with suspected PCD for diagnostic testing	1	2	3	4	5
Managing high volume of patients with confirmed PCD	1	2	3	4	5
Quality control through service assessment (e.g. ERN or government accreditation)	1	2	3	4	5
Having a multidisciplinary team	1	2	3	4	5
Appropriate setting for clinical appointment and diagnostic testing	1	2	3	4	5

35. **When providing good quality care for PCD patients,** how important are the following items. **Please rate** each of the items from 1 (not important at all) to 5 (very important).

1: Not at all important; 2: Slightly important; 3: Neutral; 4: Important; 5: Very important

Keeping up-to-date with latest findings from PCD research	1	2	3	4	5
Concentrating expertise and equipment in few centres	1	2	3	4	5
Quality control through service evaluation or assessment	1	2	3	4	5
Establishing evidence-based guidelines for PCD management	1	2	3	4	5
Having a multidisciplinary team involved in patient care	1	2	3	4	5
Engaging with patient organisations	1	2	3	4	5
Establishing international collaboration to advance PCD research	1	2	3	4	5
Operating through a PCD specialist centre	1	2	3	4	5
Increasing awareness of the disease	1	2	3	4	5
Discussing difficult cases with experts	1	2	3	4	5

36. In summary, in your country, what are the most important **barriers for delivery of care for PCD**? **Please rate** each of the items from 1 (not important at all) to 5 (very important).

1: Not at all important; 2: Slightly important; 3: Neutral; 4: Important; 5: Very important

Lack of awareness of disease	1	2	3	4	5
Limited time for clinical activities	1	2	3	4	5
Limited time for research activities	1	2	3	4	5
Limited funding	1	2	3	4	5
Difficulties for patients to reach the service	1	2	3	4	5
Lack of an adult service for PCD	1	2	3	4	5
Difficult diagnostic cases	1	2	3	4	5
Difficult management cases	1	2	3	4	5
Lack of communication between clinicians	1	2	3	4	5
Lack of standardised diagnostic testing and reporting	1	2	3	4	5
Lack of appropriate setting (e.g. room for clinics/lab) to diagnose and manage patients	1	2	3	4	5

End of survey

Thank you for taking the time to complete this survey. If you have any questions or additional comments, please contact Bruna Rubbo (b.rubbo@soton.ac.uk).

E.3 Typology for service delivery studies

Table 24. Summary of typologies for country classification

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
Organisation for Economic Co-operation and Development (OECD) (1987) (367)	Beveridge model	UK, Nordic countries, Southern European countries, Ireland	Well established classification	Old, Few categories, Only based on finance	12	13
	Bismarck model	Austria, Belgium, France, Germany, Luxembourg, Netherland				
	Private insurance	US				

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
World Health Organisation (WHO) (1997) (368)	Bismarck model with well-established financing systems	Austria, Belgium, France, Germany, Luxembourg, Netherland, Switzerland	Includes all European countries	Old, Only based on finance	16	18
	Bismarck model but in a state of transition	Israel, Turkey				
	Beveridge model with well-established financing systems	Denmark, Finland, Iceland, Ireland, Norway, Sweden, UK				
	Transitioning from insurance-based to predominantly tax-based	Greece, Italy, Portugal, Spain				

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Semashko model	Albania, Bosnia, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia, Macedonia, Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan				
Moran (2000) (369)	Entrenched command and control states	UK, Denmark, Sweden, Norway, Finland	Each category is described in detail	Based on specific	8	9

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Supply states	US		countries, generalizability ? Would need to classify other countries		
	Corporativist state	Germany				
	Insecure command and control states	Portugal, Spain, Italy, Greece				
Wendt (2009) (370)	Ideal type state healthcare system	Denmark, UK, Sweden, Norway	Detailed classification	Identified 27 different systems, too many categories	7	10
	State-based mixed type (regulation state, financing societal, provision state)	Eastern Europe				
	State-based mixed type (regulation societal, financing state, provision state)	Portugal, Greece				

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Societal-based mixed-type (regulation societal, financing societal, provision private)	Germany				
	Private-based mixed-type (regulation private, financing state, provision private)	US				
	Pure mixed-type (regulation state, financing societal, provision private)	Austria				
Wendt (2009) (371)	Health service provision-oriented type	Austria, Belgium, France, Germany, Luxembourg	Clustering analysis		10	10

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Universal coverage-controlled access type	Denmark, UK, Sweden, Italy, Ireland		Only European countries and not all of them		
	Low budget-restricted access type	Portugal, Spain, Finland				
Bohm (2013) (372)	National Health Service	Denmark, Finland, Iceland, Norway, Sweden, Portugal, Spain, UK	Reduced previous 27 categories from Wendt to 5	Assumes hierarchy (regulation > financing > provision)	15	20
	National Health Insurance System	Australia, Canada, Ireland, New Zealand, Italy				
	Social Health Insurance System	Austria, Germany, Luxembourg, Switzerland				
	Private Health System	US				

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Etatist Social Health Insurance	Estonia, Czech Republic, Hungary, Poland, Slovakia, Japan, Korea, Belgium, France, Netherlands, Israel				
Ferreira (2018) (373)	Austria-Germany	Austria, Germany	Recent classification Includes most European countries	Only EU countries, would have to exclude Switzerland, Serbia, Russia, Turkey	13	14
	Central and Northern countries	Belgium, Denmark, Finland, France, Ireland, Luxembourg, Netherlands, Sweden, UK				
	Southern countries	Cyprus, Greece, Italy, Malta, Portugal, Spain				

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Eastern countries cluster A	Bulgaria, Hungary, Latvia, Lithuania, Slovakia, Romania				
	Eastern countries cluster B	Croatia, Czech Republic, Estonia, Poland, Slovenia				
Reibling (2019) (374)	Supply- and choice-oriented public systems	Australia, Austria, Belgium, Czech Republic, Germany, France, Ireland, Iceland, Luxembourg, Slovenia	Comprehensive exhaustive clustering using multiple variables	Relies on primary care data to develop clusters, therefore less relevant for PCD	14	19
	Performance- and primary care-oriented public systems	Finland, Japan, Korea, Norway, New Zealand, Portugal, Sweden				
	Regulation-oriented public systems	Canada, Denmark, Spain, Italy, Netherlands, UK				

Author (year)	Classification	Countries	Advantages	Limitations	Countries available for comparison	Countries available for whole data
	Low-supply and low-performance-oriented mixed systems	Estonia, Hungary, Poland, Slovakia				
	Supply- and performance-oriented private systems	Switzerland, US				

List of References

1. Commission Communication on Rare Diseases. *Communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Rare Diseases: Europe's challanges.* Available from: https://ec.europa.eu/health/ph_threats/non_com/docs/rare_com_en.pdf [accessed 31st May 2020].
2. Rare Diseases Task Force Working Group on Health Indicators. *Health indicators for rare diseases: conceptual framework and development of indicators from existing sources.* Available from: https://webgate.ec.europa.eu/chafea_pdb/assets/files/pdb/20082291/20082291_d04_01_oth_en_ps.pdf [accessed 31st May 2020]
3. Lucas JS, Burgess A, Mitchison HM, Moya E, Williamson M, Hogg C. Diagnosis and management of primary ciliary dyskinesia. *Arch Dis Child.* 2014;99(9):850-6.
4. Kuehni CE, Frischer T, Strippoli MP, Maurer E, Bush A, Nielsen KG, et al. Factors influencing age at diagnosis of primary ciliary dyskinesia in European children. *Eur Respir J.* 2010;36(6):1248-58.
5. Kuehni CE, Goutaki M, Rubbo, B, Lucas, J.S. Management of primary ciliary dyskinesia: current practice and future perspectives. In: Chalmers J, Polverino E, Aliberti S. (eds.). *Bronchiectasis*. European Respiratory Society Monograph; 2018. p. 282-299.
6. Rubbo B, Best S, Hirst RA, Shoemark A, Goggin P, Carr SB, et al. Clinical features and management of children with primary ciliary dyskinesia in England. *Arch Dis Child.* 2020. Available from: doi.org/10.1136/archdischild-2019-317687.
7. O'Callaghan C, Chetcuti P, Moya E. High prevalence of primary ciliary dyskinesia in a British Asian population. *Arch Dis Child.* 2010;95(1):51-2.
8. Onoufriadis A, Paff T, Antony D, Shoemark A, Micha D, Kuyt B, et al. Splice-site mutations in the axonemal outer dynein arm docking complex gene CCDC114 cause primary ciliary dyskinesia. *Am J Hum Genet.* 2013;92(1):88-98.
9. Spagnolo PdB, R.M. The challenges of clinical research in orphan diseases. In: Cottin VC, J-F. Richeldi, L. (eds). *Orphan Lung Diseases, a clinical guide to rare lung disease.* Springer-Verlag London; 2015.
10. Sommer JU, Schafer K, Omran H, Olbrich H, Wallmeier J, Blum A, et al. ENT manifestations in patients with primary ciliary dyskinesia: prevalence and significance of otorhinolaryngologic co-morbidities. *Eur Arch Otorhinolaryngol.* 2011;268(3):383-8.
11. Behan L, Dunn Galvin A, Rubbo B, Masefield S, Copeland F, Manion M, et al. Diagnosing primary ciliary dyskinesia: an international patient perspective. *Eur Respir J.* 2016;48(4):1096-107.
12. European Commission. Regulation (EC) No 141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products. *Official Journal of the European Communities*; 1999.
13. EUROPLAN. Recommendations for the development of national plans for rare diseases, guidance document. Available from: https://www.eurordis.org/content/europlan-guidance-national-plans-and-conferences [accessed 31st May 2020].
14. The Council of the European Union. Council recommendation of 8 June 2009 on an action in the field of rare diseases. Luxembourg: Official Journal of the European Union; 2009.
15. Kobbernagel HE, Buchvald FF, Haarman EG, Casaulta C, Collins SA, Hogg C, et al. Study protocol, rationale and recruitment in a European multi-centre randomized controlled trial to determine the efficacy and safety of azithromycin maintenance therapy for 6 months in primary ciliary dyskinesia. *BMC Pulm Med.* 2016;16(1):104-.

16. Werner C, Lablans M, Ataian M, Raidt J, Wallmeier J, Große-Onnebrink J, et al. An international registry for primary ciliary dyskinesia. *Eur Respir J.* 2016 Mar;47(3):849-59.
17. Behan L, Dimitrov BD, Kuehni CE, Hogg C, Carroll M, Evans HJ, et al. PICADAR: a diagnostic predictive tool for primary ciliary dyskinesia. *Eur Respir J.* 2016;47(4):1103-12.
18. Olin JT, Burns K, Carson JL, Metjian H, Atkinson JJ, Davis SD, et al. Diagnostic yield of nasal scrape biopsies in primary ciliary dyskinesia: a multicenter experience. *Pediatr Pulmonol.* 2011;46(5):483-8.
19. Jackson CL, Behan L, Collins SA, Goggin PM, Adam EC, Coles JL, et al. Accuracy of diagnostic testing in primary ciliary dyskinesia*. Eur Respir J.* 2016;47(3):837-48.
20. Heon-Klin V. European Reference networks for rare diseases: what is the conceptual framework? *Orphanet J Rare Dis.* 2017;12(1):137.
21. Chilvers MA, Rutman A, O'Callaghan C. Ciliary beat pattern is associated with specific ultrastructural defects in primary ciliary dyskinesia. J Allergy Clin Immunol. 2003;112(3):518-24.
22. Ferkol TW, Leigh MW. Ciliopathies: the central role of cilia in a spectrum of pediatric disorders. *J Pediatr.* 2012;160(3):366-71.
23. Satir P, Heuser T, Sale WS. A Structural Basis for How Motile Cilia Beat. *Bioscience.* 2014;64(12):1073-83.
24. Olcese C, Patel MP, Shoemark A, Kiviluoto S, Legendre M, Williams HJ, et al. X-linked primary ciliary dyskinesia due to mutations in the cytoplasmic axonemal dynein assembly factor PIH1D3. *Nat Commun.* 2017;8:14279.
25. Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, et al. European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *Eur Respir J.* 2017;49(1):1601090.
26. Lucas JS, Davis SD, Omran H, Shoemark A. Primary ciliary dyskinesia in the genomics age. *Lancet Respir Med.* 2020;8(2):202-16.
27. Mitchison HM, Valente EM. Motile and non-motile cilia in human pathology: from function to phenotypes. *J Pathol.* 2017;241(2):294-309.
28. Davis SD, Ferkol TW, Rosenfeld M, Lee HS, Dell SD, Sagel SD, et al. Clinical features of childhood primary ciliary dyskinesia by genotype and ultrastructural phenotype. *Am J Respir Crit Care Med.* 2015;191(3):316-24.
29. Marshall CR, Scherer SW, Zariwala MA, Lau L, Paton TA, Stockley T, et al. Whole-Exome Sequencing and Targeted Copy Number Analysis in Primary Ciliary Dyskinesia. *G3 (Bethesda).* 2015;5(8):1775-81.
30. Reish O, Slatkin M, Chapman-Shimshoni D, Elizur A, Chioza B, Castleman V, et al. Founder mutation(s) in the RSPH9 gene leading to primary ciliary dyskinesia in two inbred Bedouin families. *Ann Hum Genet.* 2010;74(2):117-25.
31. Daniels ML, Leigh MW, Davis SD, Armstrong MC, Carson JL, Hazucha M, et al. Founder mutation in RSPH4A identified in patients of Hispanic descent with primary ciliary dyskinesia. *Hum Mutat.* 2013;34(10):1352-6.
32. Zariwala MA, Leigh MW, Ceppa F, Kennedy MP, Noone PG, Carson JL, et al. Mutations of DNAI1 in primary ciliary dyskinesia: evidence of founder effect in a common mutation. *Am J Respir Crit Care Med.* 2006;174(8):858-66.
33. Lie H, Zariwala MA, Helms C, Bowcock AM, Carson JL, Brown DE, 3rd, et al. Primary ciliary dyskinesia in Amish communities. *J Pediatr.* 2010;156(6):1023-5.
34. Ferkol TW, Puffenberger EG, Lie H, Helms C, Strauss KA, Bowcock A, et al. Primary ciliary dyskinesia-causing mutations in Amish and Mennonite communities. *J Pediatr.* 2013;163(2):383-7.
35. Shoemark A, Moya E, Hirst RA, Patel MP, Robson EA, Hayward J, et al. High prevalence of CCDC103 p.His154Pro mutation causing primary ciliary dyskinesia disrupts protein oligomerisation and is associated with normal diagnostic investigations. *Thorax.* 2018;73(2):157-66.
36. Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, et al. Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. *Thorax.* 2012;67(5):433-41.
37. Olbrich H, Haffner K, Kispert A, Volkel A, Volz A, Sasmaz G, et al. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nat Genet.* 2002;30(2):143-4.

38. Noone PG, Leigh MW, Sannuti A, Minnix SL, Carson JL, Hazucha M, et al. Primary ciliary dyskinesia: diagnostic and phenotypic features. Am J Respir Crit Care Med. 2004;169(4):459-67.
39. Mullowney T, Manson D, Kim R, Stephens D, Shah V, Dell S. Primary ciliary dyskinesia and neonatal respiratory distress. *Pediatrics.* 2014;134(6):1160-6.
40. Goutaki M, Meier AB, Halbeisen FS, Lucas JS, Dell SD, Maurer E, et al. Clinical manifestations in primary ciliary dyskinesia: systematic review and meta-analysis. *Eur Respir J.* 2016;48(4):1081-95.
41. Coren ME, Meeks M, Morrison I, Buchdahl RM, Bush A. Primary ciliary dyskinesia: age at diagnosis and symptom history. *Acta Paediatr.* 2002;91(6):667-9.
42. Bequignon E, Dupuy L, Zerah-Lancner F, Bassinet L, Honore I, Legendre M, et al. Critical Evaluation of Sinonasal Disease in 64 Adults with Primary Ciliary Dyskinesia. *J Clin Med.* 2019;8(5).
43. Alanin MC. Bacteriology and treatment of infections in the upper and lower airways in patients with primary ciliary dyskinesia: adressing the paranasal sinuses. *Dan Med J.* 2017;64(5).
44. Hosie PH, Fitzgerald DA, Jaffe A, Birman CS, Rutland J, Morgan LC. Presentation of primary ciliary dyskinesia in children: 30 years' experience. *J Paediatr Child Health.* 2015;51(7):722-6.
45. Bhatt JM, Muhonen EG, Meier M, Sagel SD, Chan KH. Rhinosinusitis in Pediatric Primary Ciliary Dyskinesia: Impact of Disease. *Otolaryngol Head Neck Surg.* 2019;161(5):877-80.
46. Pifferi M, Bush A, Caramella D, Di Cicco M, Zangani M, Chinellato I, et al. Agenesis of paranasal sinuses and nasal nitric oxide in primary ciliary dyskinesia. *Eur Respir J.* 2011;37(3):566-71.
47. Majithia A, Fong J, Hariri M, Harcourt J. Hearing outcomes in children with primary ciliary dyskinesia--a longitudinal study. *Int J Pediatr Otorhinolaryngol.* 2005;69(8):1061-4.
48. Pruliere-Escabasse V, Coste A, Chauvin P, Fauroux B, Tamalet A, Garabedian EN, et al. Otologic features in children with primary ciliary dyskinesia. *Arch Otolaryngol Head Neck Surg.* 2010;136(11):1121-6.
49. Best S, Shoemark A, Rubbo B, Patel MP, Fassad MR, Dixon M, et al. Risk factors for situs defects and congenital heart disease in primary ciliary dyskinesia. *Thorax.* 2019;74(2):203-5.
50. Munro NC, Currie DC, Lindsay KS, Ryder TA, Rutman A, Dewar A, et al. Fertility in men with primary ciliary dyskinesia presenting with respiratory infection. *Thorax.* 1994;49(7):684-7.
51. Afzelius BA, Eliasson R. Male and female infertility problems in the immotile-cilia syndrome. *Eur J Respir Dis Suppl.* 1983;127:144-7.
52. Eliasson R, Mossberg B, Camner P, Afzelius BA. The immotile-cilia syndrome. A congenital ciliary abnormality as an etiologic factor in chronic airway infections and male sterility. *N Engl J Med.* 1977;297(1):1-6.
53. Vanaken GJ, Bassinet L, Boon M, Mani R, Honore I, Papon JF, et al. Infertility in an adult cohort with primary ciliary dyskinesia: phenotype-gene association. *Eur Respir J.* 2017;50(5).
54. Goutaki M, Papon JF, Boon M, Casaulta C, Eber E, Escudier E, et al. Standardised clinical data from patients with primary ciliary dyskinesia: FOLLOW-PCD. *ERJ Open Res.* 2020;6(1).
55. Goutaki M, Eich MO, Halbeisen FS, Barben J, Casaulta C, Clarenbach C, et al. The Swiss Primary Ciliary Dyskinesia registry: objectives, methods and first results. *Swiss Med Wkly.* 2019;149.
56. Verra F, Escudier E, Bignon J, Pinchon MC, Boucherat M, Bernaudin JF, et al. Inherited factors in diffuse bronchiectasis in the adult: a prospective study. *Eur Respir J.* 1991;4(8):937-44.
57. Leigh MW, Ferkol TW, Davis SD, Lee HS, Rosenfeld M, Dell SD, et al. Clinical Features and Associated Likelihood of Primary Ciliary Dyskinesia in Children and Adolescents. *Ann Am Thorac Soc.* 2016;13(8):1305-13.
58. Kuehni CE, Lucas JS. Toward an Earlier Diagnosis of Primary Ciliary Dyskinesia. Which Patients Should Undergo Detailed Diagnostic Testing? *Ann Am Thorac Soc.* 2016;13(8):1239-43.
59. Boon M, Jorissen M, Proesmans M, De Boeck K. Primary ciliary dyskinesia, an orphan disease. *Eur J Pediatr.* 2013;172(2):151-62.

60. Dehlink E, Hogg C, Carr SB, Bush A. Clinical phenotype and current diagnostic criteria for primary ciliary dyskinesia. *Expert Rev Respir Med.* 2016;10(11):1163-75.
61. Rubbo B, Behan L, Lima de Queiroz AP, Goggin P, Jackson C, Packham S, et al. Ten years' experience of Primary Ciliary Dyskinesia Diagnostic Testing. *Eur Respir J.* 2017;50(suppl 61):PA1850.
62. Lin AE, Krikov S, Riehle-Colarusso T, Frias JL, Belmont J, Anderka M, et al. Laterality defects in the national birth defects prevention study (1998-2007): birth prevalence and descriptive epidemiology. *Am J Med Genet A.* 2014;164a(10):2581-91.
63. Rubbo B, Best S, Hirst RA, Shoemark A, Goggin P, Carr SB, et al. Clinical features and management of children with primary ciliary dyskinesia in England. *Arch Dis Child.* 2020.
64. Humbert M, Wagner TO. Rare respiratory diseases are ready for primetime: from Rare Disease Day to the European Reference Networks. *Eur Respir J.* 2017;49(2).
65. Shapiro AJ, Davis SD, Polineni D, Manion M, Rosenfeld M, Dell SD, et al. Diagnosis of Primary Ciliary Dyskinesia. An Official American Thoracic Society Clinical Practice Guideline. *Am J Respir Crit Care Med.* 2018;197(12):e24-e39.
66. Collins SA, Gove K, Walker W, Lucas JSA. Nasal nitric oxide screening for primary ciliary dyskinesia: systematic review and meta-analysis. *Eur Respir J.* 2014;44(6):1589.
67. Leigh MW, Hazucha MJ, Chawla KK, Baker BR, Shapiro AJ, Brown DE, et al. Standardizing nasal nitric oxide measurement as a test for primary ciliary dyskinesia. *Ann Am Thorac Soc.* 2013;10(6):574-81.
68. Kharitonov SA, Rajakulasingam K, O'Connor B, Durham SR, Barnes PJ. Nasal nitric oxide is increased in patients with asthma and allergic rhinitis and may be modulated by nasal glucocorticoids. *J Allergy Clin Immunol.* 1997;99(1 Pt 1):58-64.
69. Zysman-Colman ZN, Kaspy KR, Alizadehfar R, NyKamp KR, Zariwala MA, Knowles MR, et al. Nasal Nitric Oxide in Primary Immunodeficiency and Primary Ciliary Dyskinesia: Helping to Distinguish Between Clinically Similar Diseases. *J Clin Immunol.* 2019;39(2):216-24.
70. Mateos-Corral D, Coombs R, Grasemann H, Ratjen F, Dell SD. Diagnostic value of nasal nitric oxide measured with non-velum closure techniques for children with primary ciliary dyskinesia. *J Pediatr.* 2011;159(3):420-4.
71. Liu C, Zheng M, He F, Wang X, Zhang L. Role of exhaled nasal nitric oxide in distinguishing between chronic rhinosinusitis with and without nasal polyps. *Am J Rhinol Allergy.* 2017;31(6):389-94.
72. Shapiro AJ, Dell SD, Gaston B, O'Connor M, Marozkina N, Manion M, et al. Nasal Nitric Oxide Measurement in Primary Ciliary Dyskinesia. A Technical Paper on Standardized Testing Protocols. *Ann Am Thorac Soc.* 2020;17(2):e1-e12.
73. Lundberg JO, Weitzberg E, Nordvall SL, Kuylenstierna R, Lundberg JM, Alving K. Primarily nasal origin of exhaled nitric oxide and absence in Kartagener's syndrome. *Eur Respir J.* 1994;7(8):1501-4.
74. Adams PS, Tian X, Zahid M, Khalifa O, Leatherbury L, Lo CW. Establishing normative nasal nitric oxide values in infants. *Respir Med.* 2015;109(9):1126-30.
75. Marthin JK, Philipsen MC, Rosthoj S, Nielsen KG. Infant nasal nitric oxide over time: natural evolution and impact of respiratory tract infection. *Eur Respir J.* 2018;51(6).
76. Lucas JS, Walker WT. NO way! Nasal nitric oxide measurement in infants. *Eur Respir J.* 2018;51(6).
77. Collins SA, Behan L, Harris A, Gove K, Lucas JS. The dangers of widespread nitric oxide screening for primary ciliary dyskinesia. *Thorax.* 2016;71(6):560-1.
78. Knowles MR, Ostrowski LE, Leigh MW, Sears PR, Davis SD, Wolf WE, et al. Mutations in RSPH1 cause primary ciliary dyskinesia with a unique clinical and ciliary phenotype. *Am J Respir Crit Care Med.* 2014;189(6):707-17.
79. Chilvers MA, McKean M, Rutman A, Myint BS, Silverman M, Callaghan C. The effects of coronavirus on human nasal ciliated respiratory epithelium. *Eur Respir J.* 2001;18(6):965.
80. Kempeneers C, Seaton C, Garcia Espinosa B, Chilvers MA. Ciliary functional analysis: Beating a path towards standardization. *Pediatr Pulmonol.* 2019;54(10):1627-38.

81. Raidt J, Wallmeier J, Hjeij R, Onnebrink JG, Pennekamp P, Loges NT, et al. Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia. *Eur Respir J.* 2014;44(6):1579-88.
82. Lucas JS, Evans HJ, Haarman EG, Hirst RA, Hogg C, Jackson CL, et al. Exploring the Art of Ciliary Beating: The Benefits of High-Speed Video Analysis. *Chest.* 2017;152(6):1348-9.
83. Afzelius BA. A human syndrome caused by immotile cilia. *Science.* 1976;193(4250):317-9.
84. Boon M, Smits A, Cuppens H, Jaspers M, Proesmans M, Dupont LJ, et al. Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure. *Orphanet J Rare Dis.* 2014;9:11.
85. Shoemark A, Boon M, Brochhausen C, Bukowy-Bieryllo Z, De Santi MM, Goggin P, et al. International consensus guideline for reporting transmission electron microscopy results in the diagnosis of primary ciliary dyskinesia (BEAT PCD TEM Criteria). *Eur Respir J.* 2020;55(4).
86. Zariwala MA, Omran H, Ferkol TW. The emerging genetics of primary ciliary dyskinesia. *Proc Am Thorac Soc.* 2011;8(5):430-3.
87. Hirst RA, Jackson CL, Coles JL, Williams G, Rutman A, Goggin PM, et al. Culture of primary ciliary dyskinesia epithelial cells at air-liquid interface can alter ciliary phenotype but remains a robust and informative diagnostic aid. *PLoS One.* 2014;9(2):e89675.
88. Shoemark A, Frost E, Dixon M, Ollosson S, Kilpin K, Patel M, et al. Accuracy of Immunofluorescence in the Diagnosis of Primary Ciliary Dyskinesia. *Am J Respir Crit Care Med.* 2017;196(1):94-101.
89. Shoemark A, Burgoyne T, Kwan R, Dixon M, Patel MP, Rogers AV, et al. Primary ciliary dyskinesia with normal ultrastructure: three-dimensional tomography detects absence of DNAH11. *Eur Respir J.* 2018;51(2).
90. Olbrich H, Schmidts M, Werner C, Onoufriadis A, Loges NT, Raidt J, et al. Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *Am J Hum Genet.* 2012;91(4):672-84.
91. Rubbo B, Lucas JS. Clinical care for primary ciliary dyskinesia: current challenges and future directions. *Eur Respir Review.* 2017;26(145):170023.
92. Shoemark A, Dell S, Shapiro A, Lucas JS. ERS and ATS diagnostic guidelines for primary ciliary dyskinesia: similarities and differences in approach to diagnosis. *Eur Respir J.* 2019;54(3).
93. Rumman N, Jackson C, Collins S, Goggin P, Coles J, Lucas JS. Diagnosis of primary ciliary dyskinesia: potential options for resource-limited countries. *Eur Respir Rev.* 2017;26(143).
94. Werner C, Lablans M, Ataian M, Raidt J, Wallmeier J, Grosse-Onnebrink J, et al. An international registry for primary ciliary dyskinesia. *Eur Respir J.* 2016;47(3):849-59.
95. Goutaki M, Maurer E, Halbeisen FS, Amirav I, Barbato A, Behan L, et al. The international primary ciliary dyskinesia cohort (iPCD Cohort): methods and first results. *Eur Respir J.* 2017;49(1).
96. Lucas JS, Leigh MW. Diagnosis of primary ciliary dyskinesia: searching for a gold standard. *Eur Respir J.* 2014;44(6):1418-22.
97. Kempeneers C, Seaton C, Chilvers MA. Variation of Ciliary Beat Pattern in Three Different Beating Planes in Healthy Subjects. *Chest.* 2017;151(5):993-1001.
98. Papon JF, Bassinet L, Cariou-Patron G, Zerah-Lancner F, Vojtek AM, Blanchon S, et al. Quantitative analysis of ciliary beating in primary ciliary dyskinesia: a pilot study. *Orphanet J Rare Dis.* 2012;7:78.
99. Bottier M, Blanchon S, Pelle G, Bequignon E, Isabey D, Coste A, et al. A new index for characterizing micro-bead motion in a flow induced by ciliary beating: Part I, experimental analysis. *PLoS Comput Biol.* 2017;13(7):e1005605.
100. Burgoyne T, Lewis A, Dewar A, Luther P, Hogg C, Shoemark A, et al. Characterizing the ultrastructure of primary ciliary dyskinesia transposition defect using electron tomography. *Cytoskeleton (Hoboken).* 2014;71(5):294-301.
101. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American

College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-24.
102. Barbato A, Frischer T, Kuehni CE, Snijders D, Azevedo I, Baktai G, et al. Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children. *Eur Respir J.* 2009;34(6):1264.
103. Shapiro AJ, Zariwala MA, Ferkol T, Davis SD, Sagel SD, Dell SD, et al. Diagnosis, monitoring, and treatment of primary ciliary dyskinesia: PCD foundation consensus recommendations based on state of the art review. *Pediatr Pulmonol.* 2016;51(2):115-32.
104. Lucas JS, Carroll M. Primary ciliary dyskinesia and cystic fibrosis: different diseases require different treatment. *Chest.* 2014;145(4):674-6.
105. Cohen-Cymberknoh M, Simanovsky N, Hiller N, Hillel AG, Shoseyov D, Kerem E. Differences in disease expression between primary ciliary dyskinesia and cystic fibrosis with and without pancreatic insufficiency. *Chest.* 2014;145(4):738-44.
106. Paff T, Daniels JM, Weersink EJ, Lutter R, Vonk Noordegraaf A, Haarman EG. A randomised controlled trial on the effect of inhaled hypertonic saline on quality of life in primary ciliary dyskinesia. *Eur Respir J.* 2017;49(2).
107. Kuehni CE, Goutaki M, Kobbernagel HE. Hypertonic saline in patients with primary ciliary dyskinesia: on the road to evidence-based treatment for a rare lung disease. *Eur Respir J.* 2017;49(2).
108. Koh YY, Park Y, Jeong JH, Kim CK, Min YG, Chi JG. The effect of regular salbutamol on lung function and bronchial responsiveness in patients with primary ciliary dyskinesia. *Chest.* 2000;117(2):427-33.
109. Bryson HM, Sorkin EM. Dornase alfa. A review of its pharmacological properties and therapeutic potential in cystic fibrosis. *Drugs.* 1994;48(6):894-906.
110. Lucas JS, Gahleitner F, Amorim A, Boon M, Brown P, Constant C, et al. Pulmonary exacerbations in patients with primary ciliary dyskinesia: an expert consensus definition for use in clinical trials. *ERJ Open Res.* 2019;5(1).
111. Maglione M, Bush A, Montella S, Mollica C, Manna A, Esposito A, et al. Progression of lung disease in primary ciliary dyskinesia: is spirometry less accurate than CT? *Pediatr Pulmonol.* 2012;47(5):498-504.
112. Green K, Buchvald FF, Marthin JK, Hanel B, Gustafsson PM, Nielsen KG. Ventilation inhomogeneity in children with primary ciliary dyskinesia. *Thorax.* 2012;67(1):49-53.
113. Halbeisen FS, Jose A, de Jong C, Nyilas S, Latzin P, Kuehni CE, et al. Spirometric indices in primary ciliary dyskinesia: systematic review and meta-analysis. *ERJ Open Res.* 2019;5(2).
114. Halbeisen FS, Goutaki M, Spycher BD, Amirav I, Behan L, Boon M, et al. Lung function in patients with primary ciliary dyskinesia: an iPCD Cohort study. *Eur Respir J.* 2018;52(2).
115. Nyilas S, Schlegtendal A, Yammine S, Casaulta C, Latzin P, Koerner-Rettberg C. Further evidence for an association between LCI and FEV1 in patients with PCD. *Thorax.* 2015;70(9):896.
116. Tiddens H, Puderbach M, Venegas JG, Ratjen F, Donaldson SH, Davis SD, et al. Novel outcome measures for clinical trials in cystic fibrosis. *Pediatr Pulmonol.* 2015;50(3):302-15.
117. Boon M, Vermeulen FL, Gysemans W, Proesmans M, Jorissen M, De Boeck K. Lung structure-function correlation in patients with primary ciliary dyskinesia. *Thorax.* 2015;70(4):339-45.
118. Irving SJ, Ives A, Davies G, Donovan J, Edey AJ, Gill SS, et al. Lung clearance index and high-resolution computed tomography scores in primary ciliary dyskinesia. *Am J Respir Crit Care Med.* 2013;188(5):545-9.
119. Kennedy MP, Noone PG, Leigh MW, Zariwala MA, Minnix SL, Knowles MR, et al. High-resolution CT of patients with primary ciliary dyskinesia. *AJR Am J Roentgenol.* 2007;188(5):1232-8.
120. Marthin JK, Petersen N, Skovgaard LT, Nielsen KG. Lung function in patients with primary ciliary dyskinesia: a cross-sectional and 3-decade longitudinal study. *Am J Respir Crit Care Med.* 2010;181(11):1262-8.
121. Shah A, Shoemark A, MacNeill SJ, Bhaludin B, Rogers A, Bilton D, et al. A longitudinal study characterising a large adult primary ciliary dyskinesia population. *Eur Respir J.* 2016;48(2):441-50.

122. Frija-Masson J, Bassinet L, Honore I, Dufeu N, Housset B, Coste A, et al. Clinical characteristics, functional respiratory decline and follow-up in adult patients with primary ciliary dyskinesia. *Thorax.* 2017;72(2):154-60.
123. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J.* 2005;26(2):319.
124. Ellerman A, Bisgaard H. Longitudinal study of lung function in a cohort of primary ciliary dyskinesia. *Eur Respir J.* 1997;10(10):2376-9.
125. Sunther M, Bush A, Hogg C, McCann L, Carr SB. Recovery of baseline lung function after pulmonary exacerbation in children with primary ciliary dyskinesia. *Pediatr Pulmonol.* 2016;51(12):1362-6.
126. Ratjen F, Waters V, Klingel M, McDonald N, Dell S, Leahy TR, et al. Changes in airway inflammation during pulmonary exacerbations in patients with cystic fibrosis and primary ciliary dyskinesia. *Eur Respir J.* 2016;47(3):829-36.
127. Merelle ME, Schouten JP, Gerritsen J, Dankert-Roelse JE. Influence of neonatal screening and centralized treatment on long-term clinical outcome and survival of CF patients. *Eur Respir J.* 2001;18(2):306-15.
128. Comeau AM, Parad RB, Dorkin HL, Dovey M, Gerstle R, Haver K, et al. Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. *Pediatrics.* 2004;113(6):1573-81.
129. Polineni D, Davis SD, Dell SD. Treatment recommendations in Primary Ciliary Dyskinesia. *Paediatr Respir Rev.* 2016;18:39-45.
130. Alanin MC, Nielsen KG, von Buchwald C, Skov M, Aanaes K, Hoiby N, et al. A longitudinal study of lung bacterial pathogens in patients with primary ciliary dyskinesia. *Clin Microbiol Infect.* 2015;21(12):1093.e1-7.
131. Campbell RG, Birman CS, Morgan L. Management of otitis media with effusion in children with primary ciliary dyskinesia: a literature review. *Int J Pediatr Otorhinolaryngol.* 2009;73(12):1630-8.
132. Mener DJ, Lin SY, Ishman SL, Boss EF. Treatment and outcomes of chronic rhinosinusitis in children with primary ciliary dyskinesia: where is the evidence? A qualitative systematic review. *Int Forum Allergy Rhinol.* 2013;3(12):986-91.
133. Hadfield PJ, Rowe-Jones JM, Bush A, Mackay IS. Treatment of otitis media with effusion in children with primary ciliary dyskinesia. *Clin Otolaryngol Allied Sci.* 1997;22(4):302-6.
134. Parsons DS, Greene BA. A treatment for primary ciliary dyskinesia: efficacy of functional endoscopic sinus surgery. *Laryngoscope.* 1993;103(11 Pt 1):1269-72.
135. Brunner JPRCAMED. A systematic review of the treatment of chronic rhinosinusitis in adults with primary ciliary dyskinesia. *Sinusitis.* 2017.
136. Wolter NE, Dell SD, James AL, Campisi P. Middle ear ventilation in children with primary ciliary dyskinesia. *Int J Pediatr Otorhinolaryngol.* 2012;76(11):1565-8.
137. Bequignon E, Dupuy L, Escabasse V, Zerah-Lancner F, Bassinet L, Honore I, et al. Follow-Up and Management of Chronic Rhinosinusitis in Adults with Primary Ciliary Dyskinesia: Review and Experience of Our Reference Centers. *J Clin Med.* 2019;8(9).
138. Alanin MC, Aanaes K, Hoiby N, Pressler T, Skov M, Nielsen KG, et al. Sinus surgery can improve quality of life, lung infections, and lung function in patients with primary ciliary dyskinesia. *Int Forum Allergy Rhinol.* 2017;7(3):240-7.
139. Kennedy MP, Omran H, Leigh MW, Dell S, Morgan L, Molina PL, et al. Congenital heart disease and other heterotaxic defects in a large cohort of patients with primary ciliary dyskinesia. *Circulation.* 2007;115(22):2814-21.
140. Anderson RH, Devine WA, Uemura H. Diagnosis of heterotaxy syndrome. *Circulation.* 1995;91(3):906-8.

141. Van Praagh R. Terminology of congenital heart disease. Glossary and commentary. *Circulation.* 1977;56(2):139-43.
142. Peeters H, Devriendt K. Human laterality disorders. *Eur J Med Genet.* 2006;49(5):349-62.
143. Jacobs JP, Anderson RH, Weinberg PM, Walters HL, Tchervenkov CI, Del Duca D, et al. The nomenclature, definition and classification of cardiac structures in the setting of heterotaxy. *Cardiol Young.* 2007;17 Suppl 2:1-28.
144. Kim SJ, Kim WH, Lim HG, Lee JY. Outcome of 200 patients after an extracardiac Fontan procedure. *J Thorac Cardiovasc Surg.* 2008;136(1):108-16.
145. Swisher M, Jonas R, Tian X, Lee ES, Lo CW, Leatherbury L. Increased postoperative and respiratory complications in patients with congenital heart disease associated with heterotaxy. *J Thorac Cardiovasc Surg.* 2011;141(3):637-44, 44.e1-3.
146. Nakhleh N, Francis R, Giese RA, Tian X, Li Y, Zariwala MA, et al. High prevalence of respiratory ciliary dysfunction in congenital heart disease patients with heterotaxy. *Circulation.* 2012;125(18):2232-42.
147. Sironen A, Shoemark A, Patel M, Loebinger MR, Mitchison HM. Sperm defects in primary ciliary dyskinesia and related causes of male infertility. *Cell Mol Life Sci.* 2020;77(11):2029-2048.
148. Yuan S, Liu Y, Peng H, Tang C, Hennig GW, Wang Z, et al. Motile cilia of the male reproductive system require miR-34/miR-449 for development and function to generate luminal turbulence. *Proc Natl Acad Sci.* 2019;116(9):3584-93.
149. Terre B, Lewis M, Gil-Gomez G, Han Z, Lu H, Aguilera M, et al. Defects in efferent duct multiciliogenesis underlie male infertility in GEMC1-, MCIDAS- or CCNO-deficient mice. *Development.* 2019;146(8).
150. Raidt J, Werner C, Menchen T, Dougherty GW, Olbrich H, Loges NT, et al. Ciliary function and motor protein composition of human fallopian tubes. *Hum Reprod.* 2015;30(12):2871-80.
151. Quon BS, Rowe SM. New and emerging targeted therapies for cystic fibrosis. *BMJ.* 2016;352:i859.
152. Ramsey BW, Davies J, McElvaney NG, Tullis E, Bell SC, Drevinek P, et al. A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N Engl J Med.* 2011;365(18):1663-72.
153. Davies JC, Wainwright CE, Canny GJ, Chilvers MA, Howenstine MS, Munck A, et al. Efficacy and safety of ivacaftor in patients aged 6 to 11 years with cystic fibrosis with a G551D mutation. *Am J Respir Crit Care Med.* 2013;187(11):1219-25.
154. Alton E, Armstrong DK, Ashby D, Bayfield KJ, Bilton D, Bloomfield EV, et al. Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial. *Lancet Respir Med.* 2015;3(9):684-91.
155. Leigh MW, Pittman JE, Carson JL, Ferkol TW, Dell SD, Davis SD, et al. Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome. *Genet Med.* 2009;11(7):473-87.
156. Lucas JS, Chetcuti P, Copeland F, Hogg C, Kenny T, Moya E, et al. Overcoming challenges in the management of primary ciliary dyskinesia: the UK model. *Paediatr Respir Rev.* 2014;15(2):142-5.
157. Nagra A, McGinnity PM, Davis N, Salmon AP. Implementing transition: Ready Steady Go. *Arch Dis Child Educ Pract Ed.* 2015;100(6):313-20.
158. Office for National Statistics. Ethnic group, national identity and religion. Measuring equality: a guide for the collection and classification of ethnic group, national identity and religion data in the UK. Available from: https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion [accessed 31st May 2020].
159. Shapiro A.J. CMA, Davis S.D., Leigh M.W. Nasal nitric oxide and ciliary videomicroscopy: tests used for diagnosing primary ciliary dyskinesia. In: Davis S EE, Koumbourlis A. (eds). New York, NY: Springers Science and Business Media; 2015.
160. Bryant M, Ashton L, Brown J, Jebb S, Wright J, Roberts K, et al. Systematic review to identify and appraise outcome measures used to evaluate childhood obesity treatment interventions (CoOR): evidence of purpose, application, validity, reliability and sensitivity. *Health Technol Assess.* 2014;18(51):1-380.

161. Kuehni CE, Lucas JS. Diagnosis of primary ciliary dyskinesia: summary of the ERS Task Force report. *Breathe.* 2017;13(3):166-78.
162. Rubbo B, Shoemark A, Jackson CL, Hirst R, Thompson J, Hayes J, et al. Accuracy of high-speed video analysis to diagnose primary ciliary dyskinesia. *Chest.* 2019;155(5):1008-17.
163. Daniel JN. *Sampling based on the nature of the sampling unit and mixed-methods sample design.* Thousands Oaks, California: SAGE Publications, Inc; 2011.
164. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bulletin.* 1971;76(5):378-82.
165. Reichenheim ME. Confidence Intervals for the Kappa Statistic. *Stata J.* 2004;4(4):421-8.
166. Klar N, Lipsitz SR, Parzen M, Leong T. An Exact Bootstrap Confidence Interval for κ in Small Samples. *J Royal Statist Soc Series D*. 2002;51(4):467-78.
167. Genders TS, Spronk S, Stijnen T, Steyerberg EW, Lesaffre E, Hunink MG. Methods for calculating sensitivity and specificity of clustered data: a tutorial. *Radiology.* 2012;265(3):910-6.
168. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open.* 2016;6(11):e012799.
169. Shapiro AJ, Leigh MW. Value of transmission electron microscopy for primary ciliary dyskinesia diagnosis in the era of molecular medicine: Genetic defects with normal and non-diagnostic ciliary ultrastructure. *Ultrastruct Pathol.* 2017;41(6):373-85.
170. Peabody JE, Shei RJ, Bermingham BM, Phillips SE, Turner B, Rowe SM, et al. Seeing cilia: imaging modalities for ciliary motion and clinical connections. *Am J Physiol Lung Cell Mol Physiol.* 2018;314(6):L909-l21.
171. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990;43(6):543-9.
172. Gwet KL. Kappa statistics is not satisfactory for assessing the extent of agreement between raters. Available from: http://agreestat.com/papers/papers/kappa_statistic_is_not_satisfactory.pdf [accessed 31st May 2020]
173. Liechti-Gallati S, Bonsall I, Malik N, Schneider V, Kraemer LG, Ruedeberg A, et al. Genotype/phenotype association in cystic fibrosis: analyses of the delta F508, R553X, and 3905insT mutations. *Pediatr Res.* 1992;32(2):175-8.
174. Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D, Levison H, et al. The relation between genotype and phenotype in cystic fibrosis--analysis of the most common mutation (delta F508). *N Engl J Med.* 1990;323(22):1517-22.
175. Schibler A, Bolt I, Gallati S, Schoni MH, Kraemer R. High morbidity and mortality in cystic fibrosis patients compound heterozygous for 3905insT and deltaF508. *Eur Respir J.* 2001;17(6):1181-6.
176. Boon M, Wallmeier J, Ma L, Loges NT, Jaspers M, Olbrich H, et al. MCIDAS mutations result in a mucociliary clearance disorder with reduced generation of multiple motile cilia. *Nat Commun.* 2014;5:4418.
177. Wallmeier J, Al-Mutairi DA, Chen CT, Loges NT, Pennekamp P, Menchen T, et al. Mutations in CCNO result in congenital mucociliary clearance disorder with reduced generation of multiple motile cilia. *Nat Genet.* 2014;46(6):646-51.
178. Blanchon S, Legendre M, Bottier M, Tamalet A, Montantin G, Collot N, et al. Deep phenotyping, including quantitative ciliary beating parameters, and extensive genotyping in primary ciliary dyskinesia. *J Med Genet.* 2020;57(4):237-44.
179. Kispert A, Petry M, Olbrich H, Volz A, Ketelsen UP, Horvath J, et al. Genotype-phenotype correlations in PCD patients carrying DNAH5 mutations. *Thorax.* 2003;58(6):552-4.
180. Davis SD, Rosenfeld M, Lee HS, Ferkol TW, Sagel SD, Dell SD, et al. Primary Ciliary Dyskinesia: Longitudinal Study of Lung Disease by Ultrastructure Defect and Genotype. *Am J Respir Crit Care Med.* 2019;199(2):190-8.

181. Beydon N, Davis SD, Lombardi E, Allen JL, Arets HG, Aurora P, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *Am J Respir Crit Care Med.* 2007;175(12):1304-45.
182. Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW. Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease. *Am J Respir Crit Care Med.* 2013;188(8):913-22.
183. Chazal F, Michel B. An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists. 2017. Available from: https://arxiv.org/abs/1710.04019
184. Carlsson G. Topology and data. *Bull Amer Math Soc.* 2009;46:255-308.
185. Lum PY, Singh G, Lehman A, Ishkanov T, Vejdemo-Johansson M, Alagappan M, et al. Extracting insights from the shape of complex data using topology. *Sci Rep.* 2013;3:1236.
186. Singh G, Memoli, F, Zwicker M. Topological methods for the analysis of high dimensional data sets and 3D object recognition. *Eurographics Symposium on Point-Based Graphics.* The Eurographics Association; 2007.
187. Nicolau M, Levine AJ, Carlsson G. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proc Natl Acad Sci*. 2011;108(17):7265-70.
188. Nielson JL, Cooper SR, Yue JK, Sorani MD, Inoue T, Yuh EL, et al. Uncovering precision phenotype-biomarker associations in traumatic brain injury using topological data analysis. *PLoS One.* 2017;12(3):e0169490.
189. Chan JM, Carlsson G, Rabadan R. Topology of viral evolution. *Proc Natl Acad Sci*. 2013;110(46):18566-71.
190. Saul LK, Roweis ST. An introduction to Locally Linear Embedding. 2000. Available from: https://cs.nyu.edu/~roweis/lle/papers/lleintro.pdf [accessed 31st May 2020].
191. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J.* 2012;40(6):1324-43.
192. Irving S, Dixon M, Fassad MR, Frost E, Hayward J, Kilpin K, et al. Primary Ciliary Dyskinesia Due to Microtubular Defects is Associated with Worse Lung Clearance Index. *Lung.* 2018;196(2):231-8.
193. Glushakov S KI, Rekalo A. Handling missing data in clinical trials using topological data analysis. 2018.
194. Rubbo B, Behan L, Dehlink E, Goutaki M, Hogg C, Kouis P, et al. Proceedings of the COST action BM1407 inaugural conference BEAT-PCD: translational research in primary ciliary dyskinesia - bench, bedside, and population perspectives. *BMC Proc.* 2016;10(Suppl 9):66.
195. Valentgas P DNA, Wu A.W. Outcome Definition and Measurement. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. (eds). *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide.* Rockville (MD): Agency for Healthcare Research and Quality (US) Agency for Healthcare Research and Quality.; 2013.
196. Mayer-Hamblett N, Ramsey BW, Kronmal RA. Advancing outcome measures for the new era of drug development in cystic fibrosis. *Proc Am Thorac Soc.* 2007;4(4):370-7.
197. VanDevanter DR, Konstan MW. Outcome measures for clinical trials assessing treatment of cystic fibrosis lung disease. *Clin Investig*. 2012;2(2):163-75.
198. Lucas JS, Behan L, Dunn Galvin A, Alpern A, Morris AM, Carroll MP, et al. A quality-of-life measure for adults with primary ciliary dyskinesia: QOL-PCD. *Eur Respir J.* 2015;46(2):375-83.
199. Behan L, Rubbo B, Lucas JS, Dunn Galvin A. The patient's experience of primary ciliary dyskinesia: a systematic review. *Qual Life Res.* 2017;26(9):2265-85.
200. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol.* 2018;18(1):143.
201. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-73.
202. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8(1):19-32.

203. Henry RL, Mellis CM, Petrovic L. Mucoid Pseudomonas aeruginosa is a marker of poor survival in cystic fibrosis. *Pediatr Pulmonol.* 1992;12(3):158-61.
204. Cohen-Cymberknoh M, Weigert N, Gileles-Hillel A, Breuer O, Simanovsky N, Boon M, et al. Clinical impact of Pseudomonas aeruginosa colonization in patients with Primary Ciliary Dyskinesia. *Respir Med.* 2017;131:241-6.
205. Halbeisen F, Hogg C, Alanin MC, Bukowy-Bieryllo Z, Dasi F, Duncan J, et al. Proceedings of the 2nd BEAT-PCD conference and 3rd PCD training school: part 1. *BMC Proc.* 2018;12(Suppl 2):1.
206. Farley H, Rubbo B, Bukowy-Bieryllo Z, Fassad M, Goutaki M, Harman K, et al. Proceedings of the 3rd BEAT-PCD Conference and 4th PCD Training School. *BMC Proc.* 2018;12(Suppl 16):64.
207. Maglione M, Montella S, Mirra V, Bruzzese D, Santamaria F. Long-term assessment of quality of life in primary ciliary dyskinesia: time for new tools? *Chest.* 2014;146(6):e232-e3.
208. Maglione M, Bush A, Nielsen KG, Hogg C, Montella S, Marthin JK, et al. Multicenter analysis of body mass index, lung function, and sputum microbiology in primary ciliary dyskinesia. *Pediatr Pulmonol.* 2014;49(12):1243-50.
209. Joensen O, Paff T, Haarman EG, Skovgaard IM, Jensen PO, Bjarnsholt T, et al. Exhaled breath analysis using electronic nose in cystic fibrosis and primary ciliary dyskinesia patients with chronic pulmonary infections. *PLoS One.* 2014;9(12):e115584.
210. Olveira C, Padilla A, Martinez-Garcia MA, de la Rosa D, Giron RM, Vendrell M, et al. Etiology of Bronchiectasis in a Cohort of 2047 Patients. An Analysis of the Spanish Historical Bronchiectasis Registry. *Arch Bronconeumol.* 2017;53(7):366-74.
211. Mirra V, Caffarelli C, Maglione M, Valentino R, Perruolo G, Mazzarella C, et al. Hypovitaminosis D: a novel finding in primary ciliary dyskinesia. *Ital J Pediatr.* 2015;41:14.
212. Shoemark A, Wilson R. Bronchial and peripheral airway nitric oxide in primary ciliary dyskinesia and bronchiectasis. *Respir Med.* 2009;103(5):700-6.
213. Phillips GE, Thomas S, Heather S, Bush A. Airway response of children with primary ciliary dyskinesia to exercise and beta2-agonist challenge. *Eur Respir J.* 1998;11(6):1389-91.
214. Pifferi M, Bush A, Michelucci A, Di Cicco M, Piras M, Caramella D, et al. Mannose-binding lectin 2 gene polymorphism and lung damage in primary ciliary dyskinesia. *Pediatr Pulmonol.* 2015;50(2):179-86.
215. Lopes AJ, Camilo GB, de Menezes SL, Guimaraes FS. Impact of different etiologies of bronchiectasis on the pulmonary function tests. *Clin Med Res.* 2015;13(1):12-9.
216. Jain K, Padley SP, Goldstraw EJ, Kidd SJ, Hogg C, Biggart E, et al. Primary ciliary dyskinesia in the paediatric population: range and severity of radiological findings in a cohort of patients receiving tertiary care. *Clin Radiol.* 2007;62(10):986-93.
217. Chrispin AR, Norman AP. The systematic evaluation of the chest radiograph in cystic fibrosis. *Pediatr Radiol.* 1974;2(2):101-5.
218. Brody AS, Kosorok MR, Li Z, Broderick LS, Foster JL, Laxova A, et al. Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. *J Thorac Imaging.* 2006;21(1):14-21.
219. Bhalla M, Turcios N, Aponte V, Jenkins M, Leitman BS, McCauley DI, et al. Cystic fibrosis: scoring system with thin-section CT. *Radiology.* 1991;179(3):783-8.
220. Helbich TH, Heinz-Peer G, Eichler I, Wunderbaldinger P, Gotz M, Wojnarowski C, et al. Cystic fibrosis: CT assessment of lung involvement in children and adults. *Radiology.* 1999;213(2):537-44.
221. Magnin ML, Cros P, Beydon N, Mahloul M, Tamalet A, Escudier E, et al. Longitudinal lung function and structural changes in children with primary ciliary dyskinesia. *Pediatr Pulmonol.* 2012;47(8):816-25.
222. Roden L, Gorlich D, Omran H, Peters G, Grosse-Onnebrink J, Kahl BC. A retrospective analysis of the pathogens in the airways of patients with primary ciliary dyskinesia. *Respir Med.* 2019;156:69-77.

223. Rogers GB, Carroll MP, Zain NM, Bruce KD, Lock K, Walker W, et al. Complexity, temporal stability, and clinical correlates of airway bacterial community composition in primary ciliary dyskinesia. *J Clin Microbiol.* 2013;51(12):4029-35.
224. Dell SD, Leigh MW, Lucas JS, Ferkol TW, Knowles MR, Alpern A, et al. Primary Ciliary Dyskinesia: First Health-related Quality-of-Life Measures for Pediatric Patients. *Ann Am Thorac Soc.* 2016;13(10):1726-35.
225. Behan L, Leigh MW, Dell SD, Dunn Galvin A, Quittner AL, Lucas JS. Validation of a health-related quality of life instrument for primary ciliary dyskinesia (QOL-PCD). *Thorax.* 2017;72(9):832-9.
226. Rabin HR, Butler SM, Wohl ME, Geller DE, Colin AA, Schidlow DV, et al. Pulmonary exacerbations in cystic fibrosis. *Pediatr Pulmonol.* 2004;37(5):400-6.
227. Santamaria F, Montella S, Tiddens H, Guidi G, Casotti V, Maglione M, et al. Structural and functional lung disease in primary ciliary dyskinesia. *Chest.* 2008;134(2):351-7.
228. Pifferi M, Bush A, Pioggia G, Caramella D, Tartarisco G, Di Cicco M, et al. Evaluation of pulmonary disease using static lung volumes in primary ciliary dyskinesia. *Thorax.* 2012;67(11):993-9.
229. Montella S, Santamaria F, Salvatore M, Maglione M, Iacotucci P, De Santi MM, et al. Lung disease assessment in primary ciliary dyskinesia: a comparison between chest high-field magnetic resonance imaging and high-resolution computed tomography findings. *Ital J Pediatr.* 2009;35(1):24.
230. Irving S, Carr S, Hogg C, Loebinger M, Shoemark A, Bush A. Lung Clearance Index (LCI) is Stable in Most Primary Ciliary Dyskinesia (PCD) Patients Managed in a Specialist Centre: a Pilot Study. *Lung.* 2017;195(4):441-3.
231. Green K, Ejlertsen JS, Madsen A, Buchvald FF, Kongstad T, Kobbernagel H, et al. Abbreviation modalities of nitrogen multiple-breath washout tests in school children with obstructed lung disease. *Pediatr Pulmonol.* 2016;51(6):624-32.
232. Gokdemir Y, Karadag-Saygi E, Erdem E, Bayindir O, Ersu R, Karadag B, et al. Comparison of conventional pulmonary rehabilitation and high-frequency chest wall oscillation in primary ciliary dyskinesia. *Pediatr Pulmonol.* 2014;49(6):611-6.
233. Noone PG, Bennett WD, Regnis JA, Zeman KL, Carson JL, King M, et al. Effect of aerosolized uridine-5'-triphosphate on airway clearance with cough in patients with primary ciliary dyskinesia. *Am J Respir Crit Care Med.* 1999;160(1):144-9.
234. Davies JC, Alton EW. Monitoring respiratory disease severity in cystic fibrosis. *Respir Care.* 2009;54(5):606-17.
235. Nyilas S, Bauman G, Pusterla O, Sommer G, Singer F, Stranzinger E, et al. Structural and Functional Lung Impairment in Primary Ciliary Dyskinesia. Assessment with Magnetic Resonance Imaging and Multiple Breath Washout in Comparison to Spirometry. *Ann Am Thorac Soc.* 2018;15(12):1434-42.
236. Kobbernagel HE, Green K, Ring AM, Buchvald FF, Rosthøj S, Gustafsson PM, et al. One-year evolution and variability in multiple-breath washout indices in children and young adults with primary ciliary dyskinesia. *Eur Clin Respir J.* 2019;6(1):1591841-.
237. Ahmad F, Irving S, Alton E, Davies JC, Macleod K, Rosenthal M, et al. Multiple breath washouts in children can be shortened without compromising quality. *Eur Respir J.* 2015;46(6):1814-6.
238. Nyilas S, Schlegtendal A, Singer F, Goutaki M, Kuehni CE, Casaulta C, et al. Alternative inert gas washout outcomes in patients with primary ciliary dyskinesia. *Eur Respir J.* 2017;49(1).
239. Robinson P, Morgan L. Bronchiectasis in PCD looks different to CF on CT scan. *Multidiscip Respir Med.* 2018;13(Suppl 1):24.
240. Dettmer S, Ringshausen F, Vogel-Claussen J, Fuge J, Faschkami A, Shin HO, et al. Computed tomography in adult patients with primary ciliary dyskinesia: Typical imaging findings. *PLoS One.* 2018;13(2):e0191457.
241. Tadd K, Morgan L, Rosenow T, Schultz A, Susanto C, Murray C, et al. CF derived scoring systems do not fully describe the range of structural changes seen on CT scans in PCD. *Pediatr Pulmonol.* 2019;54(4):471-7.

242. Hoang-Thi TN, Revel MP, Burgel PR, Bassinet L, Honore I, Hua-Huy T, et al. Automated computed tomographic scoring of lung disease in adults with primary ciliary dyskinesia. *BMC Pulm Med.* 2018;18(1):194.
243. Campbell-Washburn AE, Ramasawmy R, Restivo MC, Bhattacharya I, Basar B, Herzka DA, et al. Opportunities in Interventional and Diagnostic Imaging by Using High-Performance Low-Field-Strength MRI. *Radiology.* 2019;293(2):384-93.
244. Edelman RR, Hatabu H, Tadamura E, Li W, Prasad PV. Noninvasive assessment of regional ventilation in the human lung using oxygen-enhanced magnetic resonance imaging. *Nat Med.* 1996;2(11):1236-9.
245. Goutaki M, Halbeisen FS, Kuehni CE. The time is right for an international PCD disease registry: insight and ongoing research activities. *Eur Respir J.* 2017;49(6).
246. Amirav I, Roduta Roberts M, Mussaffi H, Mandelberg A, Roth Y, Abitbul R, et al. Collecting clinical data in primary ciliary dyskinesia- challenges and opportunities. *F1000Res.* 2016;5:2031.
247. Kuo W, Ciet P, Tiddens HA, Zhang W, Guillerman RP, van Straten M. Monitoring cystic fibrosis lung disease by computed tomography. Radiation risk in perspective. *Am J Respir Crit Care Med.* 2014;189(11):1328-36.
248. Kubo T, Lin PJ, Stiller W, Takahashi M, Kauczor HU, Ohno Y, et al. Radiation dose reduction in chest CT: a review. *AJR Am J Roentgenol.* 2008;190(2):335-43.
249. Ioannou P, Kouis P, Kakkoura MG, Kaliva M, Toliopoulou A, Andreou K, et al. Health related quality of life in adult primary Ciliary dyskinesia patients in Cyprus: development and validation of the Greek version of the QOL-PCD questionnaire. *Health Qual Life Outcomes.* 2020;18(1):105.
250. Queiroz APLd, Athanazio RA, Olm MAK, Rubbo B, Casal YR, Lucas J, et al. Translation of the quality-of-life measure for adults with primary ciliary dyskinesia and its application in patients in Brazil. *J Bras Pneumol.* 2019;45(3):e20170358-e.
251. Neill C, Thouvenin G, Behan L, Habouria D, Corvol H, Clement A, et al. French translation and linguistic validation of the QOL-PCD, a quality of life questionnaire for patients with primary ciliary dyskinesia. *Eur Respir J.* 2016;48(suppl 60):PA4162.
252. Yılmaz Ö, Akgün M. Primary Ciliary Dyskinesia: Ready for Quality of Life Assessment. *Turk Thorac J.* 2017;18(2):27-8.
253. Reula Martin A, Behan L, Escribano A, Dasí F, Lucas J, Armengot M. Spanish version of the Quality of Life questionnaires for patients with Primary Ciliary Dyskinesia. *Eur Respir J.* 2017;50(suppl 61):PA1856.
254. Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. *Chest.* 2003;123(1):20-7.
255. Mogayzel PJ, Jr., Dunitz J, Marrow LC, Hazle LA. Improving chronic care delivery and outcomes: the impact of the cystic fibrosis Care Center Network. *BMJ Qual Saf.* 2014;23 Suppl 1:i3-8.
256. Strippoli MP, Frischer T, Barbato A, Snijders D, Maurer E, Lucas JS, et al. Management of primary ciliary dyskinesia in European children: recommendations and clinical practice. *Eur Respir J.* 2012;39(6):1482-91.
257. United Nations. Methodology: standard country or area codes for statistical use 1999. Available from: https://unstats.un.org/unsd/methodology/m49/ [accessed 31st May 2020].
258. Kouis P, Yiallouros PK, Middleton N, Evans JS, Kyriacou K, Papatheodorou SI. Prevalence of primary ciliary dyskinesia in consecutive referrals of suspect cases and the transmission electron microscopy detection rate: a systematic review and meta-analysis. *Pediatr Res.* 2017;81(3):398-405.
259. Abitbul R, Amirav I, Blau H, Alkrinawi S, Aviram M, Shoseyov D, et al. Primary ciliary dyskinesia in Israel: Prevalence, clinical features, current diagnosis and management practices. *Respir Med.* 2016;119:41-7.
260. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* 2006;3(2):77-101.

261. Boyatzis RE. Developing themes and codes. In: Boyatzis RE. (ed). *Transforming qualitative information.* Case Western Reserve University, USA: SAGE Publications Ltd; 1998. p. 29-53.
262. Fontana A, Frey, J.H. The interview: from structured questions to negotitated text. In: Denzin NK, Lincoln, Y.S. (eds). *Handbook of qualitative research.* Thousand Oaks, California: SAGE Publications, Inc.
263. Silverman D. *Qualitative Research: Theory, Method and Practice*. SAGE Publications; 2004.
264. United Nations. Methodology: Standard country or area codes for statistical use (M49). Available from: https://unstats.un.org/unsd/methodology/m49/ [accessed 31st May 2020].
265. Barbato A, Frischer T, Kuehni CE, Snijders D, Azevedo I, Baktai G, et al. Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children. *Eur Respir J.* 2009;34(6):1264-76.
266. Kobbernagel HE, Buchvald FF, Haarman EG, Casaulta C, Collins SE, et al. Efficacy and Safety of Azithromycin Maintenance Therapy in Primary Ciliary Dyskinesia (BESTCILIA): A Multicentre, Double-Blind, Randomised, Placebo-Controlled Phase 3 Trial. *Lancet Respir Med.* 2020;8:493-505.
267. Kouis P, Papatheodorou SI, Middleton N, Giallouros G, Kyriacou K, Cohen JT, et al. Cost-effectiveness analysis of three algorithms for diagnosing primary ciliary dyskinesia: a simulation study. *Orphanet J Rare Dis.* 2019;14(1):142.
268. Padman R, McColley SA, Miller DP, Konstan MW, Morgan WJ, Schechter MS, et al. Infant care patterns at epidemiologic study of cystic fibrosis sites that achieve superior childhood lung function. *Pediatrics.* 2007;119(3):e531-7.
269. Schechter MS, Fink AK, Homa K, Goss CH. The Cystic Fibrosis Foundation Patient Registry as a tool for use in quality improvement. *BMJ Qual Saf.* 2014;23 Suppl 1:i9-14.
270. Horani A, Ferkol TW, Dutcher SK, Brody SL. Genetics and biology of primary ciliary dyskinesia. *Paediatr Respir Rev.* 2016;18:18-24.
271. Werner C, Onnebrink JG, Omran H. Diagnosis and management of primary ciliary dyskinesia. *Cilia.* 2015;4(1):2.
272. Loges NT, Antony D, Maver A, Deardorff MA, Gulec EY, Gezdirici A, et al. Recessive DNAH9 Loss-of-Function Mutations Cause Laterality Defects and Subtle Respiratory Ciliary-Beating Defects. *Am J Hum Genet.* 2018;103(6):995-1008.
273. Mitchison HM, Schmidts M, Loges NT, Freshour J, Dritsoula A, Hirst RA, et al. Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nature genetics.* 2012;44(4):381-S2.
274. Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, et al. DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nature genetics.* 2013;45(9):995-1003.
275. Hoben IM, Hjeij R, Olbrich H, Dougherty GW, Nothe-Menchen T, Aprea I, et al. Mutations in C11orf70 Cause Primary Ciliary Dyskinesia with Randomization of Left/Right Body Asymmetry Due to Defects of Outer and Inner Dynein Arms. *Am J Hum Genet.* 2018;102(5):973-84.
276. Guo T, Tan ZP, Chen HM, Zheng DY, Liu L, Huang XG, et al. An effective combination of whole-exome sequencing and runs of homozygosity for the diagnosis of primary ciliary dyskinesia in consanguineous families. *Sci Rep.* 2017;7(1):7905.
277. Nunez-Olle M, Jung C, Terre B, Balsiger NA, Plata C, Roset R, et al. Constitutive Cyclin O deficiency results in penetrant hydrocephalus, impaired growth and infertility. *Oncotarget.* 2017;8(59):99261-73.
278. Alanin MC, Nielsen KG, Von Buchwald C, Skov M, Aanaes K, Høiby N, et al. A longitudinal study of lung bacterial pathogens in patients with primary ciliary dyskinesia. *Clin Microbiol Infect.* 2015;21(12):1093.e1-.e7.
279. Boon M, Smits A, Cuppens H, Jaspers M, Proesmans M, Dupont LJ, et al. Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure. *Orphanet J Rare Dis.* 2014;9(1):11.
280. Boon M, Vermeulen FL, Gysemans W, Proesmans M, Jorissen M, De Boeck K. Lung structure-function correlation in patients with primary ciliary dyskinesia. *Thorax.* 2015;70(4):339-45.

281. Bush A, Payne D, Pike S, Jenkins G, Henke MO, Rubin BK. Mucus properties in children with primary ciliary dyskinesia: comparison with cystic fibrosis. *Chest.* 2006;129(1):118-23.
282. Carotenuto M, Esposito M, Di Pasquale F, De Stefano S, Santamaria F. Psychological, cognitive and maternal stress assessment in children with primary ciliary dyskinesia. *World J Pediatr.* 2013;9(4):312-7.
283. Cockx M, Gouwy M, Godding V, De Boeck K, Van Damme J, Boon M, et al. Neutrophils from Patients with Primary Ciliary Dyskinesia Display Reduced Chemotaxis to CXCR2 Ligands. *Front Immunol.* 2017;8:1126.
284. Cohen-Cymberknoh M, Weigert N, Gileles-Hillel A, Breuer O, Simanovsky N, Boon M, et al. Clinical impact of Pseudomonas aeruginosa colonization in patients with Primary Ciliary Dyskinesia. *Respir Med.* 2017;131:241-6.
285. Davis SD, Ferkol TW, Rosenfeld M, Lee H-S, Dell SD, Sagel SD, et al. Clinical Features of Childhood Primary Ciliary Dyskinesia by Genotype and Ultrastructural Phenotype. *Am J Respir Crit Care Med.* 2015;191(3):316-24.
286. Hellinckx J, Demedts M, De Boeck K. Primary ciliary dyskinesia: evolution of pulmonary function. *Eur J Pediatr.* 1998;157(5):422-6.
287. Bush A, Chodhari R, Collins N, Copeland F, Hall P, Harcourt J, et al. Primary ciliary dyskinesia: current state of the art. *Arch Dis Child.* 2007;92(12):1136-40.
288. Kawakami M, Hattori Y, Nakamura S. Reflection of structural abnormality in the axoneme of respiratory cilia in the clinical features of immotile cilia syndrome. *Intern Med.* 1996;35(8):617-23.
289. Kennedy MP, Noone PG, Carson J, Molina PL, Ghio A, Zariwala MA, et al. Calcium stone lithoptysis in primary ciliary dyskinesia. *Respir Med.* 2007;101(1):76-83.
290. Li AM, Sonnappa S, Lex C, Wong E, Zacharasiewicz A, Bush A, et al. Non-CF bronchiectasis: does knowing the aetiology lead to changes in management? *Eur Respir J.* 2005;26(1):8-14.
291. Loomba RS, Danduran M, Nielsen KG, Ring AM, Kovach J, Anderson RH. Cardiopulmonary Exercise Testing in Fontan Patients With and Without Isomerism (Heterotaxy) as Compared to Patients With Primary Ciliary Dyskinesia and Subjects With Structurally Normal Hearts. *Pediatr Cardiol.* 2017;38(2):410-7.
292. Madsen A, Green K, Buchvald F, Hanel B, Nielsen KG. Aerobic fitness in children and young adults with primary ciliary dyskinesia. *PLoS One.* 2013;8(8):e71409.
293. Maglione M, Bush A, Montella S, Mollica C, Manna A, Esposito A, et al. Progression of lung disease in primary ciliary dyskinesia: Is spirometry less accurate than CT? *Pediatr Pulmonol.* 2012;47(5):498-504.
294. Maglione M, Montella S, Mirra V, Bruzzese D, Santamaria F. Long-term Assessment of Quality of Life in Primary Ciliary Dyskinesia. *Chest.* 2014;146(6):e232-e3.
295. Maglione M, Montella S, Mollica C, Carnovale V, Iacotucci P, De Gregorio F, et al. Lung structure and function similarities between primary ciliary dyskinesia and mild cystic fibrosis: a pilot study. *Ital J Pediatr.* 2017;43(1):34.
296. Schluchter MD, Konstan MW, Drumm ML, Yankaskas JR, Knowles MR. Classifying severity of cystic fibrosis lung disease using longitudinal pulmonary function data. *Am J Respir Crit Care Med.* 2006;174(7):780-6.
297. McManus IC, Mitchison HM, Chung EM, Stubbings GF, Martin N. Primary ciliary dyskinesia (Siewert's/Kartagener's syndrome): respiratory symptoms and psycho-social impact. *BMC Pulm Med.* 2003;3:4.
298. McManus IC, Stubbings GF, Martin N. Stigmatization, physical illness and mental health in primary ciliary dyskinesia. *J Health Psychol.* 2006;11(3):467-82.
299. Mirra V, Caffarelli C, Maglione M, Valentino R, Perruolo G, Mazzarella C, et al. Hypovitaminosis D: a novel finding in primary ciliary dyskinesia. *Ital J Pediatr.* 2015;41(1):14.
300. Montella S, Santamaria F, Salvatore M, Pignata C, Maglione M, Iacotucci P, et al. Assessment of chest high-field magnetic resonance imaging in children and young adults with

noncystic fibrosis chronic lung disease: comparison to high-resolution computed tomography and correlation with pulmonary function. *Invest Radiol.* 2009;44(9):532-8.
301. Montuschi P, Paris D, Montella S, Melck D, Mirra V, Santini G, et al. Nuclear magnetic resonance-based metabolomics discriminates primary ciliary dyskinesia from cystic fibrosis. *Am J Respir Crit Care Med.* 2014;190(2):229-33.
302. Stern RC. The diagnosis of cystic fibrosis. *N Engl J Med.* 1997;336(7):487-91.
303. Bush A, Hogg C. Primary ciliary dyskinesia: recent advances in epidemiology, diagnosis, management and relationship with the expanding spectrum of ciliopathy. *Expert Rev Respir Med.* 2012;6(6):663-82.
304. Nyilas S, Schlegtendal A, Singer F, Goutaki M, Kuehni CE, Casaulta C, et al. Alternative inert gas washout outcomes in patients with primary ciliary dyskinesia. *Eur Respir J.* 2017;49(1):1600466.
305. Oktem S, Karadag B, Erdem E, Gokdemir Y, Karakoc F, Dagli E, et al. Sleep disordered breathing in patients with primary ciliary dyskinesia. *Pediatr Pulmonol.* 2013;48(9):897-903.
306. Paff T, van der Schee MP, Daniels JM, Pals G, Postmus PE, Sterk PJ, et al. Exhaled molecular profiles in the assessment of cystic fibrosis and primary ciliary dyskinesia. *J Cyst Fibros.* 2013;12(5):454-60.
307. Pifferi M, Bush A, Di Cicco M, Pradal U, Ragazzo V, Macchia P, et al. Health-related quality of life and unmet needs in patients with primary ciliary dyskinesia. *Eur Respir J.* 2010;35(4):787-94.
308. Santamaria F, Esposito M, Montella S, Cantone E, Mollica C, De Stefano S, et al. Sleep disordered breathing and airway disease in primary ciliary dyskinesia. *Respirology.* 2014;19(4):570-5.
309. Shah A, Shoemark A, Macneill SJ, Bhaludin B, Rogers A, Bilton D, et al. A longitudinal study characterising a large adult primary ciliary dyskinesia population. *Eur Respir J.* 2016;48(2):441-50.
310. Smit HJ, Schreurs AJ, Van den Bosch JM, Westermann CJ. Is resection of bronchiectasis beneficial in patients with primary ciliary dyskinesia? *Chest.* 1996;109(6):1541-4.
311. Svobodova T, Djakow J, Zemkova D, Cipra A, Pohunek P, Lebl J. Impaired Growth during Childhood in Patients with Primary Ciliary Dyskinesia. *Int J Endocrinol.* 2013;2013:731423.
312. Tamalet A, Clement A, Roudot-Thoraval F, Desmarquest P, Roger G, Boule M, et al. Abnormal central complex is a marker of severity in the presence of partial ciliary defect. *Pediatrics.* 2001;108(5):E86.
313. Valerio G, Giallauria F, Montella S, Vaino N, Vigorito C, Mirra V, et al. Cardiopulmonary assessment in primary ciliary dyskinesia. *Eur J Clin Invest.* 2012;42(6):617-22.
314. Vallet C, Escudier E, Roudot-Thoraval F, Blanchon S, Fauroux B, Beydon N, et al. Primary ciliary dyskinesia presentation in 60 children according to ciliary ultrastructure. *Eur J Pediatr.* 2013;172(8):1053-60.
315. Wells GD, Wilkes DL, Schneiderman JE, Rayner T, Elmi M, Selvadurai H, et al. Skeletal muscle metabolism in cystic fibrosis and primary ciliary dyskinesia. *Pediatr Res.* 2011;69(1):40-5.
316. Whalley S, McManus IC. Living with primary ciliary dyskinesia: a prospective qualitative study of knowledge sharing, symptom concealment, embarrassment, mistrust, and stigma. *BMC Pulm Med.* 2006;6:25.
317. Yiallouros PK, Kouis P, Middleton N, Nearchou M, Adamidi T, Georgiou A, et al. Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients. *Respir Med.* 2015;109(3):347-56.
318. Zihlif N, Paraskakis E, Lex C, Van de Pohl LA, Bush A. Correlation between cough frequency and airway inflammation in children with primary ciliary dyskinesia. *Pediatr Pulmonol.* 2005;39(6):551-7.
319. Zihlif N, Paraskakis E, Tripoli C, Lex C, Bush A. Markers of airway inflammation in primary ciliary dyskinesia studied using exhaled breath condensate. *Pediatr Pulmonol.* 2006;41(6):509-14.
320. Horsley AR, Gustafsson PM, Macleod KA, Saunders C, Greening AP, Porteous DJ, et al. Lung clearance index is a sensitive, repeatable and practical measure of airways disease in adults with cystic fibrosis. *Thorax.* 2008;63(2):135-40.

321. Lee TW, Brownlee KG, Conway SP, Denton M, Littlewood JM. Evaluation of a new definition for chronic Pseudomonas aeruginosa infection in cystic fibrosis patients. *J Cyst Fibros.* 2003;2(1):29-34.
322. Lucas JS, Behan L, Dunn Galvin A, Alpern A, Morris AM, Carroll MP, et al. A quality-of-life measure for adults with primary ciliary dyskinesia: QOL–PCD. *Eur Respir J.* 2015;46(2):375-83.
323. Behan L, Leigh MW, Dell SD, Quittner AL, Hogg C, Lucas JS. Validation of pediatric health-related quality of life instruments for primary ciliary dyskinesia (QOL-PCD). *Pediatr Pulmonol.* 2019.
324. Roelants M, Hauspie R, Hoppenbrouwers K. References for growth and pubertal development from birth to 21 years in Flanders, Belgium. Ann Hum Biol. 2009;36(6):680-94.
325. Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, et al. Consensus statement for inert gas washout measurement using multiple- and single- breath tests. *Eur Respir J.* 2013;41(3):507-22.
326. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J.* 2005;26(2):319-38.
327. Wang X, Dockery DW, Wypij D, Fay ME, Ferris BG, Jr. Pulmonary function between 6 and 18 years of age. *Pediatr Pulmonol.* 1993;15(2):75-88.
328. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med.* 1999;159(1):179-87.
329. Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, et al. Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med.* 2008;177(3):253-60.
330. Jones M, Castile R, Davis S, Kisling J, Filbrun D, Flucke R, et al. Forced expiratory flows and volumes in infants. Normative data and lung growth. *Am J Respir Crit Care Med.* 2000;161(2 Pt 1):353-9.
331. Davis SD, Rosenfeld M, Kerby GS, Brumback L, Kloster MH, Acton JD, et al. Multicenter evaluation of infant lung function tests as cystic fibrosis clinical trial endpoints. *Am J Respir Crit Care Med.* 2010;182(11):1387-97.
332. Rosenthal M, Bain SH, Cramer D, Helms P, Denison D, Bush A, et al. Lung function in white children aged 4 to 19 years: I--Spirometry. *Thorax.* 1993;48(8):794-802.
333. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. *Eur Respir J.* 2005;26(5):948-68.
334. Zapletal A, Samanek M, Paul T. Upstream and total airway conductance in children and adolescents. *Bull Eur Physiopathol Respir.* 1982;18(1):31-7.
335. Standardized lung function testing. Report working party. *Bull Eur Physiopathol Respir.* 1983;19 Suppl 5:1-95.
336. Duiverman EJ, Clement J, van de Woestijne KP, Neijens HJ, van den Bergh AC, Kerrebijn KF. Forced oscillation technique. Reference values for resistance and reactance over a frequency spectrum of 2-26 Hz in healthy children aged 2.3-12.5 years. *Bull Eur Physiopathol Respir.* 1985;21(2):171-8.
337. Bibi H, Montgomery M, Pasterkamp H, Chernick V. Relationship between response to inhaled salbutamol and methacholine bronchial provocation in children with suspected asthma. *Pediatr Pulmonol.* 1991;10(4):244-8.
338. Benden C. The Chrispin-Norman score in cystic fibrosis: doing away with the lateral view. *Eur Respir J.* 2005;26(5):894-7.
339. Knowles MR, Ostrowski LE, Leigh MW, Sears PR, Davis SD, Wolf WE, et al. Mutations in RSPH1 cause primary ciliary dyskinesia with a unique clinical and ciliary phenotype. *Am J Respir Crit Care Med.* 2014;189(6):707-17.
340. Chai H, Farr RS, Froehlich LA, Mathison DA, McLean JA, Rosenthal RR, et al. Standardization of bronchial inhalation challenge procedures. J Allergy Clin Immunol. 1975;56(4):323-7.
341. Godfrey S. Physiological response to exercise in children with lung or heart disease. *Arch Dis Child.* 1970;45(242):534-8.

342. Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest.* 1984;85(6):751-8.
343. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. *J Bras Pneumol.* 2007;33(4):397-406.
344. Neder JA, Andreoni S, Castelo-Filho A, Nery LE. Reference values for lung function tests. I. Static volumes. *Braz J Med Biol Res.* 1999;32(6):703-17.
345. Koopman M, Zanen P, Kruitwagen CL, van der Ent CK, Arets HG. Reference values for paediatric pulmonary function testing: The Utrecht dataset. *Respir Med.* 2011;105(1):15-23.
346. Kirkby J, Stanojevic S, Welsh L, Lum S, Badier M, Beardsmore C, et al. Reference equations for specific airway resistance in children: the Asthma UK initiative. *Eur Respir J.* 2010;36(3):622-9.
347. Cotes JE, Chinn DJ, Quanjer PH, Roca J, Yernault JC. Standardization of the measurement of transfer factor (diffusing capacity). Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl.* 1993;16:41-52.
348. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl.* 1993;16:5-40.
349. Soni R, Dobbin CJ, Milross MA, Young IH, Bye PP. Gas exchange in stable patients with moderate-to-severe lung disease from cystic fibrosis. *J Cyst Fibros.* 2008;7(4):285-91.
350. Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Stat Med.* 1998;17(4):407-29.
351. Puderbach M, Eichinger M, Gahr J, Ley S, Tuengerthal S, Schmahl A, et al. Proton MRI appearance of cystic fibrosis: comparison to CT. *Eur Radiol.* 2007;17(3):716-24.
352. Gaultier C, Boule M, Allaire Y, Clement A, Buvry A, Girard F. Determination of capillary oxygen tension in infants and children: assessment of methodology and normal values during growth. *Bull Eur Physiopathol Respir.* 1979;14(3):287-97.
353. Goldberg LR. The development of markers for the Big-Five factor structure. *Psychol Assess.* 1992;4(1):26-42.
354. Goldberg LR. Analyses of Digman's child-personality data: derivation of Big-Five factor scores from each of six samples. *J Pers.* 2001;69(5):709-43.
355. Hofstee WKB, Kiers HA, de Raad B, Goldberg LR, Ostendorf F. A comparison of Big-Five structures of personality traits in Dutch, English, and German. *Eur J Person.* 1997;11(1):15-31.
356. Bushnell DM, Martin ML. Quality of life and Parkinson's disease: translation and validation of the US Parkinson's Disease Questionnaire (PDQ-39). *Qual Life Res.* 1999;8(4):345-50.
357. Yammine S, Singer F, Gustafsson P, Latzin P. Impact of different breathing protocols on multiple-breath washout outcomes in children. *J Cystic Fibros.* 2014;13(2):190-7.
358. Naidich DP, McCauley DI, Khouri NF, Stitik FP, Siegelman SS. Computed tomography of bronchiectasis. *J Comput Assist Tomogr.* 1982;6(3):437-44.
359. Silverman M, Anderson SD. Standardization of Exercise Tests in Asthmatic Children. *Arch Dis Child.* 1972;47(256):882-9.
360. Pifferi M, Caramella D, Bulleri A, Baldi S, Peroni D, Pietrobelli A, et al. Pediatric bronchiectasis: correlation of HRCT, ventilation and perfusion scintigraphy, and pulmonary function testing. *Pediatr Pulmonol.* 2004;38(4):298-303.
361. Kraynack NC, Gothard MD, Falletta LM, McBride JT. Approach to treating cystic fibrosis pulmonary exacerbations varies widely across US CF care centers. *Pediatr Pulmonol.* 2011;46(9):870-81.
362. Gustafsson PM, De Jong PA, Tiddens HA, Lindblad A. Multiple-breath inert gas washout and spirometry versus structural lung disease in cystic fibrosis. *Thorax.* 2008;63(2):129-34.
363. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am J Respir Crit Care Med.* 2005;171(8):912-30.

364. Corey M, Levison H, Crozier D. Five- to seven-year course of pulmonary function in cystic fibrosis. *Am Rev Respir Dis.* 1976;114(6):1085-92.
365. Schneiderman-Walker J, Wilkes DL, Strug L, Lands LC, Pollock SL, Selvadurai HC, et al. Sex differences in habitual physical activity and lung function decline in children with cystic fibrosis. *J Pediatr.* 2005;147(3):321-6.
366. Wells GD, Wilkes DL, Schneiderman-Walker J, Elmi M, Tullis E, Lands LC, et al. Reliability and validity of the habitual activity estimation scale (HAES) in patients with cystic fibrosis. *Pediatr Pulmonol.* 2008;43(4):345-53.
367. The Organisation of Economic Co-operation and Development. OECD Economic Outlook. 1987, issue 11987.
368. Saltman RB, Figueras J, World Health Organization. Regional Office for E. European health care reform : analysis of current strategies. In: Saltman RB, Figueras J. (eds). Copenhagen : WHO Regional Office for Europe; 1997. Available from: https://apps.who.int/iris/handle/10665/107267 [accessed 31st May 2020].
369. Moran M. Understanding the Welfare State: The Case of Health Care. *Brit J Polit Int Rel.* 2000;2(2):135-60.
370. Wendt C, Frisina L, Rothgang H. Healthcare System Types: A Conceptual Framework for Comparison. *Soc Policy Admin.* 2009;43(1):70-90.
371. Wendt C. Mapping European healthcare systems: a comparative analysis of financing, service provision and access to healthcare. *J Eur Soc Policy.* 2009;19(5):432-45.
372. Bohm K, Schmid A, Gotze R, Landwehr C, Rothgang H. Five types of OECD healthcare systems: empirical results of a deductive classification. *Health Policy.* 2013;113(3):258-69.
373. Ferreira PL, Tavares AI, Quintal C, Santana P. EU health systems classification: a new proposal from EURO-HEALTHY. *BMC Health Serv Res.* 2018;18(1):511.
374. Reibling N, Ariaans M, Wendt C. Worlds of Healthcare: A Healthcare System Typology of OECD Countries. *Health Policy.* 2019;123(7):611-20.