Online supplementary files

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 (1-3). While systematic reviews focus on gathering evidence to address a specific question, scoping reviews map the relevant literature in the field of interest and therefore has a broader scope. Scoping reviews are useful to a) map the types of evidence that are available in a given field, particularly where this has not been comprehensively reviewed before; b) identify key concepts and definitions in the literature, highlighting inconsistencies; c) develop specific questions; and d) explore gaps in the existing literature.

We opted to perform a scoping review as it fulfilled our aims. We therefore mostly 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 and what was found throughout the manuscript.

Box 1. Key terms used in the search strategy in Embase, with results from each search term (*n* articles retrieved).

1. Exp kartagener syndrome/ (3011)
2. Exp ciliary motility disorders/ (4947)
3. primary ciliary dyskinesia.ti,ab. (3430)
4. 1 OR 2 OR 3 (6922)

5. exp respiratory function test/ or exp lung function test/ (426284)
6. exp vital capacity/ (38202)
7. exp spirometry/ (67349)
8. exp airway resistance/ (27257)
9. exp blood gas analysis/ (61158)
10. exp bronchial provocation test/ (12139)
11. capnometry/ or exp lung function test/ or patient monitoring/ (561554)
12. exp lung compliance/ (18950)
13. exp lung volume measurements/ (174399)
14. exp plethysmography, whole body/ (5794)
15. exp pulmonary gas exchange/ (33550)
16. Bronchiectasis.ti,ab. (26951)
17. exp bronchiectasis/co, di, dm, ep, et, pc, su [Complication, Diagnosis, Disease Management, Epidemiology, Etiology, Prevention, Surgery] (5957)
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 (727603)

19. Outcome parameter$.mp. or Treatment outcome/ [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading] (1823035)
20. exp hospital admission$/ or patient readmission/ or hospitalization/ (704024)
21. Hospital$.mp. (4223046)
22. mortality/ (846510)
23. morbidity/ (383659)
24. life expectancy/ (67955)
25. (Day$ antibiotic$ or antibiotic$ course$).mp. (3368)
26. Need for surgery.mp. (8952)
27. Quality of life/ (657721)
28. Disease progression (348419)
29. 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 (7289447)

30. Pulmonary exacerbation$.mp OR Disease exacerbation/ (278715)
31. Respiratory rate.mp (36742)
32. Monitoring, physiologic/ (57445)
33. Cough/ OR cough frequency.mp (56043)
34. Respiratory sounds/ OR respiratory frequency.mp OR breathing frequency.mp (25829)
35. Rhinomanometry/ OR exp Otorhinolaryngologic Surgical Procedures/ OR exp Otorhinolaryngologic Diseases/ (994984)
36. Sputum/ OR sputum clearance.mp OR sputum colo?r.mp (53716)
37. 30 or 31 or 32 or 33 or 34 or 35 or 36 (1467807)

38. Tomography, Emission-Computed/ or tomography.mp. (1919855)
39. Magnetic Resonance Imaging/ or MRI.mp. (1232707)
40. Radiography/ or Xray.mp. or Radiography.mp. (1057370)
41. Diagnostic Techniques, ontological/ OR Hearing tests/ OR Audiometry/ (52493)
42. Exp Otitis Media/ (63538)
43. Body mass index/ (499610)
44. Symptom score.mp OR symptom scale.mp (37036)
45. Inflammation/ or Inflammatory markers.mp or biomarkers/ (1134591)
46. 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 (5276241)

47. 18 or 29 or 37 or 46 (12954712)
48. 4 AND 47 (5547)
49. limit 48 to (human and yr="1996 -Current") (2145)
50. remove duplicates from 49 (2112)

Supplementary Table 1. Characteristics of included studies

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) (112)	Israel	2012 to 2013	Multicentre prospective study	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) (39)	UK	January 2008 to May 2014	Retrospective study	19	Healthy controls (17)	Median: 13.89	Not reported	Not reported
Alanin *et al* (2017) (64)	Denmark	November 2013 to February 2016	Prospective uncontrolled pre and post intervention cohort study	24	N/A	Median: 24, range: 10 to 65	Clinical phenotype + TEM, HSVA or genetic testing	Inclusion: definite PCD and above 6 years of age
Alanin *et al* (2015) (70)	Denmark	January 2002 to December 2012	Retrospective cohort	107	N/A	Median: 17, range 0 to 74	Clinical symptoms + (TEM, HSVA or genetic testing)	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**
Anagnostopoulou *et al* (2018) (40)	Switzerland, Germany	Not reported	Retrospective study	17	40	Mean: 11.8, range: 5.1–18.1	Not reported	Inclusion: free from acute respiratory disease for at least 2 weeks prior to testing. Exclusion: for healthy controls, patients with asthma or other respiratory disease, history of prematurity, and bone, neuromuscular or cardiac disease that could affect lung function were excluded.
Behan *et al* (2017) (71)	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 +	Adults (aged ≥18 years) with diagnosis of PCD in one of the specified diagnostic centres and ability to read and speak English fluently.

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 and/or genetic testing.	
Boon *et al* (2014) (113)	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
Boon *et al* (2015) (12)	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) (86)	UK	Not reported	Unclear	19	CF children (30)	Mean (SD): 9.5 (3)	nNO, CBF and TEM	Not reported
Carotenuto *et al* (2013) (72)	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

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**
								disease, mental retardation (IQ less than 70), epilepsy, and psychiatric disorders
Cockx *et al* (2017 a) (87)	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, defined as no change in cough or sputum, no fever, no change in therapy for a period of at least 2 weeks, change in forced expiratory volume in 1 second (FEV1) < 10% since the last measurement
Cockx *et al* (2017 b) (88)	Belgium	June 2012 to November 2016	Case-control	36	Health controls (numbers not reported)	Mean: 13, range 2 to 26	HSVA, cell culture, TEM, genetic testing	As above (see Cockx *et al* 2017 a)
Cohen-Cymberknoh	Israel	Not reported	Cross-sectional	20	60 patients with CF	Mean (SD): Adults 25.8	According to the ATS diagnostic	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**
et al (2019) (82)			surveys			(5.7), children 10.4 (3.5)	guidelines (17)	
Cohen-Cymberknoh *et al* (2017) (65)	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 (19)	Patients with follow-up data for at least 3 years + results from at least 2 sputum cultures
Cohen-Cymberknoh *et al* (2014) (51)	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
Davis *et al* (2015) (20)	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
Davis *et al* (2019) (21)	USA, Canada	2006 to 2011	Prospective, longitudinal,	137	N/A	Mean (SD): 7.8 (4.6)	TEM or genetic testing	<19 years at enrolment and ≥ 2 annual study

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**
			multicentre, observational study					visits and confirmed PCD
Dettmer *et al* (2018) (52)	Germany	2011 to 2017	Retrospective study	46	75 bronchiectasis patients	Median (range): 38 (18 to 72)	Patients with definite or probable PCD, according to Werner et al.	Exclusion: CT with insufficient quality due to a slice thickness >5mm or to severe motion artefacts
Eden *et al* (2019) (114)	USA	2008 to 2017	Longitudinal study	79	58 alpha-1 antitrypsin deficiency, 18 common variable immunodeficiency, 460 idiopathic	Mean (SD): 41.9 (14.5)	Characteristic clinical manifestations + genetic studies, mucosal biopsy, and nasal nitric oxide	Exclusion: Patients with CF
Ellerman *et al* (1997) (22)	Denmark	Late 1970s to 1994, with minimum of	Prospective cohort	24	N/A	Median (range): 21 (2 to 56)	Clinical phenotype + HSVA + normal	Inclusion: confirmed diagnosis + regular spirometry.

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**
		2 years follow-up					sweat test to exclude CF	Exclusion: CF patients
Emiralioglu *et al* (2020) (115)	Turkey	January 2013 to December 2018	Cohort study	46	N/A	Median age at diagnosis (range): 8.5 (6 months to 15 years)	Clinical and radiological findings, nNO, HSVA, genetic testing, TEM	15 patients (out of the original 61) were excluded due to potential novel candidate genes
Frija-Masson *et al* (2017) (116)	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
Fuger *et al* (2018) (23)	France	2000 to 2015	Cross-sectional study from retrospective cohort	42	73 CF	Median (IQR): 8.9 (6.4 to 13.5)	Typical clinical characteristics or suggestive clinical features + TEM or genetic testing	Exclusion: immunodeficiencies, diseases that could alter PaO_2, pancreatic sufficient CF
Gokdemir *et*	Turkey	Not reported	Randomised	24	N/A	Mean (SD):	Clinical	Inclusion: clinical stability

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**
al (2014) (24)			controlled crossover study			12.9 (2.7), range 7 to 18	phenotype or TEM	Exclusion: history of pneumothorax, massive hemoptysis or congestive heart failure
Goutaki *et al* (2017) (68)	Australia, Belgium, Cyprus, Denmark, France, Germany, Israel, Italy, Netherlands, Norway, Poland, Serbia, Switzerland,	Up to April 2016	Cross-section of retrospective cohort	1609	N/A	Range: 0 to 19	Clinical characteristics, nNO, HSVA, TEM, genetic testing	All patients included in the international PCD cohort study

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**
	Turkey, UK, USA, Canada							
Green *et al* (2012) (41)	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) (42)	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 (19)	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
Halbeisen *et al* (2018)	Australia, Belgium,	Up to April 2016	Cross-sectional	991	N/A	Not reported	Clinical characteristics,	Inclusion: All patients in the international PCD

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**
(25)	Cyprus, Denmark, France, Germany, Israel, Italy, Netherlands, Norway, Poland, Serbia, Switzerland, Turkey, UK		retrospective study				nNO, HSVA, TEM, genetic testing	cohort study that had data on FEV_1 and FVC. Exclusion: < 6 years, no lung function available, insufficient information to calculate z-scores
Hellinckx *et al* (1998) (26)	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

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**
Hoang-Thi *et al* (2018) (53)	France	November 2009 to July 2016	Retrospective study	62	N/A	Mean (SD): 39 (15)	According to the ESR guidelines (36)	Inclusion: CT exams performed between November 2009 and July 2016 + spirometric measurements performed within a 6-month period
Ioannou *et al* (2020) (73)	Cyprus	January 2017 to June 2019	Cross-sectional study	31	N/A	Median: 33.6	Combination of nNO, TEM, HSVA, and genetic testing	Patients with definite or highly likely diagnosis of PCD according to the ERS guidelines (36); age >18 years; and ability to speak and read Greek fluently
Irving *et al* (2013) (45)	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* (39).	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**
Irving *et al* (2017) (44)	UK	2009 to 2010; 2014 to 2015	Prospective cohort	29	N/A	Median: 14, range 3 to 53	TEM or genetic testing	Not reported
Irving *et al* (2018) (43)	UK	Not reported	Cross-sectional study	69	N/A	Median (range): 13 (4 to 41)	nNO, HSVA, TEM, genetic testing	Definite or highly likely PCD according to European guidelines (36)
Jain *et al* (2007) (54)	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
Joensen *et al* (2014) (100)	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

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**
								measurement
Kawakami *et al* (1996) (101)	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
Kenis-Coskun *et al* (2019) (74)	Turkey	May 2018 to May 2019	Cross-sectional study	19	44	Mean (SD, range): 10.31 (1.73, 7 to 13)	TEM or low nNO or dextrocardia + typical clinical findings	Exclusion: acute exacerbation or hospital admittance in the last 2 weeks
Kennedy *et al* (2007a) (102)	USA	August 2003 to March 2006 for prospective study; Prior to August 2003 for retrospective study	Mixture of prospective and retrospective study	142	N/A	Mean (SD) for *n*=7 with outcome measure: 56 (7)	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**
Kennedy *et al* (2007b) (55)	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 (46)
Knowles *et al* (2014) (117)	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 sex matched)	TEM or genetic testing	Not reported
Kobbernagel *et al* (2019) (46)	Denmark	Not reported	Single-centre, prospective, observational, longitudinal study	42	N/A	Median (range): 15.4 (6.5 to 29.7)	HSVA, TEM, genetic testing	School-aged children and young adults (aged >5 to <30 years) with a confirmed diagnosis of PCD, and clinically stable at the baseline visit
Kobbernagel	Denmark,	June 2014 to	Multicentre,	90	N/A	Range: 7 to	Clinical	Inclusion: predicted FEV_1

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**
et al (2020) (92)	Germany, Netherlands, Switzerland, UK	August 2016	double-blind, randomised, placebo-controlled phase 3 trial			50	characteristics, nNO, HSVA, TEM, IF, genetic testing	>40%; received at least 30 days of antibiotics for respiratory tract infections or exacerbations within the preceding 2 years; currently received no systemic or inhaled maintenance antibiotics; and had not taken azithromycin within 1 month before screening. Exclusion: current infection (at screening) with *Achromobacter xylosoxidans* or *Burkholderia cepacia* complex, infection with non-tuberculous

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**
								mycobacteria within 6 months, or chronic infection with *Pseudomonas aeruginosa* (defined as culture of *Pseudomonas aeruginosa* in 50% or more of the sputum samples within the last year, provided at least three sputum cultures were available). Other exclusions were: allergic reaction to macrolide antibiotics or other ingredients of the study drug; alanine transaminase twice or more the upper limit of normal or history of

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**
								portal hypertension; serum creatinine concentrations greater than 150 µmol/L or glomerular filtration rate of less than 50 mL/min; prolonged QT interval, cardiac arrhythmia, severe heart failure, or electrolyte disturbances; myasthenia gravis; treatment with medicinal products known to possibly interact with azithromycin or prolong QT interval (appendix p 1); pregnancy, breastfeeding, or fertile women using unreliable

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**
								contraception; or use of home oxygen or assisted ventilation
Koh *et al* (2000) (27)	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
Koucky *et al* (2020) (47)	Czech Republic	Not reported	Cross-sectional study	11	24 CF, 15 allergic bronchial asthma, 19 control	Median (range): 7.8 (0.6 to 15.8)	TEM	Patients with confirmed PCD diagnosis
Li *et al* (2005) (56)	UK	1986 to 2002	Retrospective study	20	N/A	Not reported	Clinical phenotype, nNO (14% of cases), LM (49% of cases), TEM (70%	Inclusion: HRCT-diagnosed bronchiectasis in subjects with suggestive clinical features

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**
							of cases)	Exclusion: CF diagnosed by sweat test and/or analysis of genetic testing.
Loomba *et al* (2017) (94)	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* (54)	Not reported
Lopes *et al* (2015) (28)	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

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.
Madsen *et al* (2013) (95)	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	Inclusion: children and young adults; healthy age-, gender- and BMI-matched non-atopic

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**
							(n=39) (19)	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) (57)	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 stable patient, and of a second CT scan plus spirometry during exacerbation. Exclusion: < 6 years of age, unable to perform

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**
								spirometry, or had only one CT scan during follow-up
Maglione *et al* (2014a) (29)	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
Maglione *et al* (2014b) (75)	Italy	Not reported	Prospective questionnaire	20	N/A	Median: 16.9; range 12 to 33.4	Not reported	Not reported
Maglione *et al* (2017) (58)	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* (60). PCD patients: stable lung disease, without 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**
								dyspnea or cough, no pulmonary function changes and no requirement for intravenous antibiotics in the previous 4 weeks. Exclusion: acute 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) (59)	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	Clinical phenotype, HSVA, TEM, computerised EM (for IDA	Inclusion: age < 15 years at the beginning of follow-up, at least 8 years of follow-up, at least 2 concurrent CT 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**
						13.8	defects, after 2002)	lung function tests available in a phase of clinical stability of the lung disease without modification of the treatment regimen in the last 4 weeks.
Marino *et al* (2019) (103)	UK	September 2016 top April 2017	Prospective study	43	N/A	Range: 0 to 16	According to the ERS guidelines (36)	Not reported
Marthin *et al* (2010) (30)	Denmark	Late 1970s onwards	Partly cross-sectional and partly designed as an uncontrolled, observational, single-group, single-centre, longitudinal	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

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**
			and retrospective study of prospectively collected data					
McManus *et al* (2003) (76)	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) (77)	UK	January 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) (104)	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* (56) Exclusion: airway infections or asthma

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**
								exacerbations during the 4 weeks prior to enrolment, current 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) (60)	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) (61)	Italy	March 2007 to June 2008	Prospective, cross-sectional study	14	Primary immunodeficiency patients (14), recurrent	Median: 15.2, range 10.4 to 29.3	Clinical phenotype, LM, TEM	Inclusion: patients with PCD, chronic lung disorders, primary immunodeficiency,

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**
					pneumonia (13)			recurrent pneumonia 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) (105)	Italy	Not reported	Cross-sectional study	45	Primary analysis: CF (21), age-matched healthy controls (21) Validation subjects: CF (25), age-matched healthy controls (25)	Mean (SD) primary Analysis: 17.4 (0.9), range 11 to 32 Mean (SD) validation subjects:	PCD and CF were diagnosed according to published criteria (70, 71)	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**
						15.7 (0.6), range 11 to 31		
Noone *et al* (1999) (106)	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) (118)	USA	1994 to 2002	Cohort study	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
Nyilas *et al* (2017) (48)	Germany and Switzerla	March 2013 to April 2015	Cross-sectional multicentre	49	37	Mean (SD): 14.7 (6.6), range 11 to	Clinical phenotype, HSVA + (TEM, IF	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**
	nd		study			18	or genetic testing)	
Nyilas *et al* (2018) (49)	Germany	April 2015 to February 2016	Prospective cross-sectional, single-centre, observational study	30	N/A	Median (range): 13.4 (5 to 28)	According to ERS consensus (76)	Absence of acute pulmonary exacerbation during the last 3 weeks before the study
Oktem *et al* (2013) (83)	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) (31)	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	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**
							diagnosis	
Paff *et al* (2013) (107)	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 (19)	Exclusion: children with any pulmonary, inflammatory or metabolic disease.
Paff *et al* (2017) (89)	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 sulphate allergy, or in use

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**
								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 agent, were not eligible to proceed in the trial.
Phillips *et al*	UK	Not reported	Cross-	12	12	Median: 11,	Clinical	Inclusion for healthy

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**
(1998) (32)			sectional			range 7 to 15	phenotype, HSVA, TEM	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.
Piatti *et al* (2020) (93)	Italy	2007 to 2017	Single-centre, retrospective, cross-sectional study	58	N/A	Children mean (range): 11.1 (2 to 17) Adults mean (range): 39.4 (19 to 70)	Cardiac situs, nNO, HSVA, TEM, genetic testing	Clinical cases of PCD that have been diagnosed and followed-up during the last 10 years
Pifferi *et al* (2010) (78)	Italy	Dec 2007 to May 2008	Cross-sectional (questionnaire	78	N/A	Mean (SD): 21.4 (12.9), range 1.7 to	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**
			s)			48.5		
Pifferi *et al* (2012) (33)	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) (119)	Italy	Not reported	Cross-sectional	45	53	Median (IQR): 14 (22.25)	nNO, (HSVA + TEM, n=37), (HSVA + cell culture, n=8)	Not reported
Pifferi *et al* (2017) (108)	Italy	Not reported	Cross-sectional and prospective longitudinal	51	35 secondary ciliary dyskinesia, 10 controls	Median (IQR): 24.5 (22.9)	HSVA, TEM, cell culture, according to ERS consensus and	All subjects aged ≥6 years, with a diagnosis of PCD. For secondary ciliary dyskinesia, PCD

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**
			(subset)				guidelines (36, 76)	was excluded as reported in Pifferi *et al* (87).
Ratjen *et al* (2016) (90)	Canada	Not reported	Cross-sectional + prospective cohort study	35	17	Median (IQR): 11.0 (6.8 to 15.3)	Clinical phenotype, nNO, TEM, genetic testing	Inclusion: at least 6 years at enrolment; ability to perform reproducible 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

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**
								haemoptysis; patients with *P. aeruginosa* or *Burkholderia cepacia* complex infection (for CF only)
Ring *et al* (2018) (96)	Denmark	Not reported	Prospective, observational, single-centre, cohort study	36	61 CF patients	Mean (range) at visit 1: 11.8 (6 to 18) Mean (range) at visit 2: 12.9 (7 to 18)	Clinical characteristics, nNO, HSVA, TEM	Inclusion: all patients with a definite diagnosis of PCD Exclusion: not able to perform exercise test or loss to follow between study visits
Roden *et al* (2019) (66)	Germany	2010 to March 2016	Cross-sectional and retrospective longitudinal study	106 cross-sectional; 28 longitudinal	N/A	Cross-sectional median (range): 13 (0 to 71) Longitudinal	According to ERS guidelines (36)	All patients with at least one respiratory specimen were included for cross-sectional analysis; all patients with at least 4 respiratory specimens for

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**
						median (range): 8 (0 to 41)		the longitudinal analysis
Rogers *et al* (2013) (67)	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) (84)	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) (84)	Italy	Not reported	Cross-sectional, prospective study	16	42	Median: 10.4, range 4.9 to 17.2	HSVA, TEM	Inclusion: lung disease stability, ability to perform reliable pulmonary function

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**
								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 psychoactive drugs; craniofacial

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**
								abnormalities, neuromuscular disorders or concomitant genetic diseases such as Trisomy 21 or Prader–Willi syndrome
Shah *et al* (2016) (34)	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) (109)	UK	March 2005 to March 2007 and January 2006 to June 2006	Case-control	20	Non-PCD bronchiectasis (20), healthy controls (20)	40 (95%CI 32-45)	LM, TEM	Inclusion for healthy controls: no history of respiratory disease and free from bacterial or viral infections for 8 weeks before study Exclusion: 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**
								smokers, CF patients (screened by sweat test, followed by CF genotyping), history of asthma
Simsek *et al* (2018) (97)	Turkey	December 2013 to March 2014	Unclear	31	29 healthy controls	Mean (SD): 13.3 (3.0)	According to the ERS guidelines (36)	Clinically stable with no change in medication for at least 3 weeks, and able to cooperate with the measurements
Sismanlar *et al* (2018) (85)	Turkey	Not reported	Case-control study	15	31 healthy controls	Mean (SD): 12.4 (0.88)	Clinical symptoms, nNO, HSVA, TEM	Exclusion: acute upper and/or lower airway infection, chronic oxygen supplementation, inability to perform pulmonary function tests, patients with other chronic diseases
Smit *et al*	The	1952 to 1994	Retrospective	21	N/A	Age at	Clinical	Exclusion: language

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**
(1996) (110)	Netherlands		cohort study			present, (range): 46 (32-61) for lung resection group (n=13); 46 (24-66) for group without lung resection (n=8)	phenotype (n=8) or TEM + HSVA (n=13)	barrier, psychiatric problems, and living abroad
Smith *et al* (2018) (50)	UK	Not reported	Multi-centre cross-sectional study	11	N/A	Mean: 13.3	Not reported	Free from pulmonary exacerbation on the day of testing and not undergoing any new acute treatments
Sunther *et al* (2016) (35)	UK	January 2003 to April 2013	Retrospective cohort study	30	N/A	Median: 11.4, range	Clinical phenotype, nNO,	Inclusion: aged 6 to 16 years, able to perform

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**
						6 to 16.2	HSVA, TEM	spirometry, history of at least one pulmonary exacerbation Exclusion: incomplete set of spirometric assessments
Svobodova *et al* (2013) (69)	Czech Republic	Not reported	Retrospective cohort study	29	N/A	Median: 14.5, range 1.5 to 24	Clinical phenotype, HSVA, TEM, genetic testing (for ODA only)	Not reported
Tadd *et al* (2019) (63)	Australia	Not reported	Multi-centre cross-sectional study	41	N/A	Mean (range): 13 (2 to 48)	According to the ATS guidelines (17)	Undergone at least 1 CT scan when clinically stable.
Tamalet *et al* (2001) (36)	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

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**
Valerio *et al* (2012) (98)	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
Valero-Moreno *et al* (2020) (79)	Spain	Not reported	Cross-sectional study	12	36 healthy controls	Mean (SD, range): 12.96 (2.71, 9 to 18)	Not reported	Not reported
Vallet *et al* (2013) (37)	France	Not reported	Retrospective study	60	N/A	Range 0 to 15	Clinical phenotype, HSVA, TEM	Not reported
Videbaek *et al* (2019)	Denmark	Not reported	Retrospective longitudinal	85	N/A	Median (range): 8.6	According to ERS guidelines and	Known genotype and at least 2 years of lung

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**
(38)			study			(4.4 to 63.6	consensus (36, 76)	function measurements
Wells *et al* (2011) (99)	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
Whalley *et al* (2006) (80)	UK	July 2005 to January 2006	Prospective qualitative interview	12	N/A	Mean: 49.8, range 27 to 65	Not reported	Inclusion: living within 250 km from London Exclusion: < 18 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**
			study (matched-pairs design)					
Yiallouros *et al* (2015) (120)	Cyprus	1998 to 2013	Cross-sectional	30	N/A	Median: 24.3, range 0.7 to 63.7	TEM + (nNO, HSVA)	Not reported
Zengin Akkus *et al* (2019) (81)	Turkey	Not reported	Cross-sectional study	14	17 CF, 15 healthy controls	Mean (SD): 46.5 (17.5)	According to the ERS consensus (76)	Exclusion: known neurologic disease
Zihlif *et al* (2005) (111)	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 baseline FEV_1 of less than 40% predicted

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**
Zihlif *et al* (2006) (91)	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.

Supplementary Table 2. Definition of outcome measures, stratified by study outcome and population descriptor

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Main study outcome: Spirometry and/or body plethysmography			
Davis *et al* (2015) (20)	**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. Spirometric measurements were expressed as percent predicted and infant lung function as z-scores. 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.
Davis *et al* (2019) (21)	**Spirometry** (FEV_1 % predicted)**, Anthropometry** (weight percentile, height percentile, BMI percentile)**, Microbiology**	None	Reported above (see Davis *et al*, 2015)

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Ellerman *et al* (1997) (22)	**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 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.
Fuger et al (2018) (23)	Capillary blood test (PaO2, PaCO2, PaO2/PaCO2 z-scores), spirometry (FEV1, FVC, FEF25-75 and FEV1/FVC z-scores, RV, RV/TLC, and TLC % predicted)	CT (presence of bronchiectasis), anthropometry (BMI z-scores), microbiology (presence of Pseudomonas aeruginosa in sputum)	Capillary blood test: Ear lobe capillary blood gas was performed and the mean of 2 to 4 capillary results was recorded.
Gokdemir *et al* (2014) (24)	**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. 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 1^{st} and 5^{th} 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.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
			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).
Halbeisen *et al* (2018) (25)	**Spirometry** (FEV_1, FVC z-scores and % predicted)	**Anthropometry** (BMI)	FEV_1 and FVC z-scores adjusted for age, sex, height and ethnicity, and % predicted values using the GLI 2012 reference values. For patients with multiple measurements, the measurement recorded at the youngest age was used. Patients under the age of 6 years were excluded to ensure better measurement quality and comparability with published CF data.
Hellinckx *et al* (1998) (26)	**Spirometry** (FEV_1, FVC, change in FEV_1 and FVC % predicted), **body plethysmography** (thoracic gas volume, total lung capacity, residual volume, and airway resistance)	**None**	Spirometry: according to ERS guidelines, the best of 3 maximal expiratory flow volume manoeuvres was analysed. All measurements were expressed as % of predicted values for sex and height according to Zapletal *et al*. Body plethysmography: single breath diffusing capacity and Krogh factor were measures according to ERS guidelines (136). FEV_1, vital capacity and Raw % predicted were calculated according to Zapletal *et al*. TLC, RV, thoracic gas volume and single breath diffusing capacity % predicted were calculated according to ERS guidelines. Reference values for total respiratory system resistance and reactance were according to Duiverman *et al*. All tests were done before and 20 min after administration of 200 µg of salbutamol. Drug dose was chosen according to Bibi *et al*.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Koh *et al* (2000) (27)	**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*. 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 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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			achieved.
Lopes *et al* (2015) (28)	**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. Spirometry/body plethysmography: All tests followed the standards formulated by the ATS (114). 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. 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* (150), 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.
Maglione *et al* (2014a) (29)	**Anthropometry** (height, weight and BMI z-scores),	**None**	Spirometry: according to published criteria. FEV_1 z score <-1.96 was considered abnormal. Anthropometry: BMI z-scores were calculated according to Cole *et al*.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	spirometry (FEV_1, FVC and FEF_{25-75} % predicted and z-scores), **microbiology**		Microbiology: chronic pseudomonal airway infection: presence of *Pseudomonas aeruginosa* for at least 6 months, with at least 3 positive cultures.
Marthin *et al* (2010) (30)	**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. 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.
Olveira *et al* (2017) (31)	**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	**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*. 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
	pathogen, by *Pseudomonas aeruginosa*, by *Haemophilus influenzae*), **treatment** (inhaled antibiotics), **CT**		
Phillips *et al* (1998) (32)	**Spirometry** (changes in % in FEV_1 and PEFR in response to exercise and to bronchodilator, baseline measurements FEV_1, FVC, FEF_{25-75} and PEFR % predicted)	**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. 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 volume loops were recorded before and 15 min after administration of the bronchodilator.
Pifferi *et al* (2012) (33)	**Body plethysmography** (FEV_1, FVC, FEF_{25-75}, FRCpleth, RV, TLC,	**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)

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	RV/TLC, airway resistance (Raw), specific airway resistance (sRaw) and effective specific resistance (sReff) % predicted and z-scores), **HRCT**		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. 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).
Shah *et al* (2016) (34)	**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	**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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	mycobacteria infection, allergic bronchopulmonary aspergillosis, other pathogens, cumulative sputum analysis)		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 (150). 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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			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.
Sunther *et al* (2016) (35)	**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) (36)	**Spirometry** (FEV_1 % predicted), **blood gas** (mean arterial PO_2)	**CT** (bronchiectasis, radiologic deterioration, lobectomy),	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.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		treatment (antibiotic use)	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 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).
Vallet *et al* (2013) (36)	**Spirometry** (FEV_1, FVC and FRC % predicted, *n* abnormal FRC and FEV_1), **blood gas** (PaO_2, *n* hypoxemic patients), **CT**	**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
	(bronchiectasis, progressive bronchiectasis)		and/or extension to a new segment). Radiological deterioration was defined as the extension of bronchiectasis.
Videbaek *et al* (2019) (38)	**Spirometry** (FEV_1, FVC, FEV_1/FVC % predicted and z-scores), **microbiology** (presence of *Pseudomonas aeruginosa*)	None	Spirometry: performed according to ATS/ERS guidelines. GLI reference equation was used to normalise spirometry parameters. Microbiology: patients were classified according to *Pseudomonas aeruginosa* infection status in 4 groups according to sputum culture results and level of precipitating antibodies (precipitins) against *Pseudomonas* using microbiology data from the latest 2 years of observation: chronic infection, intermittent infection, not-positive and not-classifiable. Chronic infection with Pseudomonas was defined as >4 samples per year with >50% positive sputum cultures and/or positive precipitins (≥ 2). Intermittent infection was defined as >4 samples per year with <50% but at least 1 positive sputum culture and negative precipitins (value 0 or 1). Patients not positive for *Pseudomonas* was defined as >4 samples per year with no positive sputum cultures and negative precipitins. Patients were deemed not classifiable if they had <4 samples per year and negative precipitins.
Main study outcome: MBW			
Ahmad *et al* (2015) (39)	**MBW** (correctly categorised %, mean time saved in	None	MBW: conducted according to published standardised protocol. Correctly categorised was defined as % of correctly predicted values using the upper limit of normal, calculated from healthy controls. Reference was 'LCI

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	seconds, mean time saved %, coefficient of variance)		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.
Anagnostopoulou *et al* (2018) (40)	**MBW** ($LCI_{standard}$, functional residual capacity (FRC), cumulative expiratory volume (CEV))	**Anthropometry** (weight z-score, height z-score)	MBW: Each child performed 3 to 4 N_2MBW according to the current consensus statement. $LCI_{standard}$ was calculated according to current recommendations, *i.e.* end-tidal nitrogen concentration (Cet) defined as the average value between 95% and 98% of expired volume and LCI as the ratio of CEV to FRC (CEV/FRC) at the first of three consecutive breaths below the cut-off of 2.5% (1/40th). Anthropometry: z-scores were calculated according to Centers for Disease Control and Prevention growth charts.
Green *et al* (2012) (41)	**SF_6 MBW** (LCI absolute values and z-scores, LCI within-session variability (CV)%, S_{acin}, S_{cond}, FRC_{SF6}), **spirometry** (FEV_1, FVC, FEF_{25-75}, FEV_1/FVC ratio z-	**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 $1/40^{th}$ 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. Abnormal lung function was defined as z-scores < -1.96. The upper limit of normal was defined as the predicted mean plus 1.96 SD for MBW variables and the lower limit of normal as

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	scores)		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) (42)	**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. GLI reference equation was used to obtain z-scores and a z score < -1.64 was considered an abnormal spirometric value.
Irving *et al* (2018) (43)	**MBW** (LCI)	**Spirometry** (FEV_1, FEF_{25-75} z-scores)**, microbiology** (presence of *Pseudomonas aeruginosa*)	Spirometry: FEV1 and FEF25–75 z-scores were calculated using the GLI. MBW: minimum of 2 runs of acceptable quality were required, in accordance with ERS/ATS guidelines. Abnormal LCI was defined as > 7.4.
Irving *et al* (2017) (44)	**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.
Irving *et al* (2013) (45)	**Spirometry** (FEV_1, FVC and MEF_{25-75} z-scores), **MBW** (LCI	**Microbiology** (infection with *Pseudomonas*	Spirometry: performed according to ATS/ERS recommendations. Subjects completed a minimum of 3 forced expiratory manoeuvres, and FEV_1 (L) and FVC (L) were expressed as z-scores.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	and functional residual capacity), **HRCT**	*aeruginosa*)	MBW: LCI was defined as the number of volume turnovers of the lungs required to reduce an inert gas to 1/40th 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.
Kobbernagel *et al* (2019) (46)	**Spirometry** (FEV_1, FCV, FEV_1/FVC z-scores and % predicted)**, MBW** (LCI, M_1/M_0, M_2/M_0, $S_{cond}*V_T$, $S_{acin}*V_T$)	**Microbiology, anthropometry** (BMI)	Respiratory exacerbation was defined as worsening respiratory symptoms at test occasion leading to the start of systemic antibiotic treatment within 1 week before or at the visit. Chronic infection by Pseudomonas aeruginosa: pathogen was cultured in ≥50% of the mucus samples from the past year, provided at least 4 annual samples were provided. N_2 MBW: performed according to the ERS/ATS consensus statement. Spirometry: measured according to ERS/ATS standards. % predicted values and z-scores were calculated using all-ages prediction equations for spirometry from the GLI. Microbiology: mucus samples were included in the data analysis if they originated

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			within 1 week before or after the test occasion, and the mucus samples were considered positive for bacteria if the culture was positive, regardless of the microscopy results.
Kouchy *et al* (2020) (47)	**MBW** ($LCI_{2.5}$, Sacin*Vt and Scond*Vt, functional residual capacity)**,** **spirometry** (FEV_1, FVC and $MMEF_{25-75}$ % predicted)**,** **endobronchial thickness** (reticular basement membrane width)**,** **bronchoalveolar lavage (fluid cytology)**	**Anthropometry** (weight, height and BMI z-scores), **microbiology**	MBW: N2-MBW adhering to relevant recommendations. Spirometry: In children aged ≥ 4 years, forced spirometry was performed according to ERS/ATS recommendations. FVC, FEV_1, and maximal mid-expiratory flow were measured and compared to the GLI 2012 reference values. Endobronchial thickness: measured using computer image analysis software according to previously validated criteria. Microbiology: *Burkholderia cepacia* complex ever positive in respiratory cultures; chronic Haemophilus influenzae, Pseudomonas aeruginosa and Staphylococcus aureus infections were defined as positive in > 50% of respiratory cultures in the last year.
Nyilas *et al* (2017) (48)	**MBW/SBW** ($LCI_{2.5\%}$, $LCI_{5\%}$, S_{acin}, S_{cond}, S_{acin}*, S_{cond}*, M1/M0,	**Microbiology** (chronic colonisation), **treatment** (use of	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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	M2/M0, SIII-DTG z-scores), **body plethysmography** (FEV_1 and FEF_{25-75} z-scores)	antibiotic long-term therapy)	consensus (166). 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.
Nyilas *et al* (2018) (49)	**Structural and functional MRI** (Eichinger score)**, MBW** (LCI, Scond and Sacin z-scores)**, spirometry** (FEV_1 and FVC z-scores)	**Anthropometry** (weight, height)	MRI: Eichinger MRI morphological score was used to assess the presence and extent of structural lung disease: 0 (not present); 1 (present and affecting 50% or less of the lobe); or 2 (present and affecting greater than 50% of the lobe). The lobe scores for each component were summed to produce a score out of 12. The total morphology score is composed of five sub-scores each with a maximum score of 12 (maximum score = 60). Functional MRI imaging, MP decomposition method was applied to generate maps of regional fractional ventilation. The distribution of ventilation and perfusion was assessed, and a threshold was applied to determine the degree of impairment. The relative fractional ventilation (RFV) and relative perfusion (RQ) impairment were calculated and expressed as a percentage of lung volume for each study participant. To estimate the degree of functional abnormalities in patients with PCD we relied on historical normal values for MRI.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			Spirometry according to current guidelines. N_2-MBW: performed in accordance with current consensus guidelines. Spirometry: a calculated z-scores from recommended reference equations for spirometry. To assess the prevalence and concordance of structural and functional outcomes abnormality was defined at ±1.64 z-scores for spirometry and MBW outcomes, structural MRI sub-scores of 2 points or greater (indicates >15% structural impairment), and functional MRI outcomes, RFV of 24.2% or greater and RQ of 19.3% or greater, according to healthy reference data.
Smith *et al* (2018) (50)	**MRI, MBW** (LCI, Scond, Sacin, ventilation defect %, coefficient of variance of ventilated image signal intensity)**, spirometry** (FEV_1, FEV_1/FVC z-scores)	**Anthropometry** (height, weight)	MRI: Three-dimensional volumetric hyperpolarized helium-3 ventilation MRI and 1H anatomical images were acquired during the same breath-hold. From these images two indices were calculated: 1) ventilation defect percentage (VDP), which quantifies the percentage of the lung volume that is not ventilated; and 2) the mean coefficient of variance of ventilated image signal intensity (CV), a metric of regional ventilation heterogeneity. 1H steady-state free precession magnetic resonance images were separately acquired for assessment of lung morphology and mucus. MBW was performed as previously described, and the parameters LCI, ventilation heterogeneity in the convection-dependent airways (Scond), and ventilation heterogeneity in diffusion–convection-dependent airways (Sacin) were calculated.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			The upper limit of normal for LCI was defined as >7.4 (119).
Main study outcome: High-resolution computed tomography			
Boon *et al* (2015) (12)	**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. The mean LCI of at least two technically acceptable measurements was used. Spirometry was performed according to the ATS/ERS guidelines. 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

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
			Flemish reference equations.
Cohen-Cymberknoh *et al* (2014) (51)	**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. FEV_1 was presented as % predicted, according to Wang *et al* for children and Hankinson *et al* 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.
Dettmer *et al* (2018) (52)	**CT** (Reiff score, lobar distribution, type of bronchiectasis, collateral findings)	**Microbiology, spirometry** (FEV_1 and FVC %predicted)**, anthropometry** (BMI)**,**	CT: Bronchiectasis was diagnosed according to the criteria described by Naidich. The Reiff-score was used to evaluate bronchiectasis. Each lobe (with the lingula considered as a separate lobe) was scored for the extent of involvement (0 = none, 1 = one or partial segment, 2 = two or more segments); severity of bronchial

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		number of exacerbations	dilatation (0 = normal, 1 = less than twice the diameter, 2 = 2–3 times the diameter, and 3 = more than 3x the diameter of the adjacent pulmonary artery); severity of the bronchial wall thickening (0 = normal, 1 = half the diameter, 2 = 0.5 to 1x diameter, and 3 = more than 1x the diameter of the adjacent pulmonary artery); type of bronchiectasis (1 = cylindrical, 2 = varicose, or 3 = cystic). The lobar distribution of bronchiectasis (0 = widespread, 1 = predominantly upper lobe, 2 = predominantly middle lobe, 3 = predominantly lower lobe, 4 = middle and lower lobes equally involved, or 5 = unclassifiable) was registered. In case of situs inversus or heterotaxy, right-sided changes were assigned to the left site according to the architecture of the lobes. collateral findings were registered. Therefore, mucous plugging, tree in bud, peripheral and central consolidations, peripheral and central ground glass opacities, interlobular septal thickening and intralobular lines were scored (0 = none, 1 = 1–3 bronchopulmonary segments involved, 2 = >3 bronchopulmonary segments involved) for the whole lung. Mosaic attenuation, atelectasis, emphysema and situs inversus / heterotaxy were classified as present / absent. It was subsequently indicated if bronchiectasis was predominant in the middle and lower lobes and if both mucous plugging and tree in bud were present in more than three segments. All terms were used according to the definition of the Fleischner Society. Subtotal or total atelectasis or a condition after resection of a lower or middle lobe or lingula was registered.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Hoang-Thi *et al* (2018) (53)	**Spirometry** (FEV_1, FVC % predicted)**, CT** (Bhalla score)	**Anthropometry** (BMI)	CT: lung structural changes were assessed by visual scoring, histogram analysis and thresholding of high attenuating lung structures. Images were scored by one thoracic radiologist using the Bhalla score. Twenty randomly selected examinations were also independently scored by a second radiologist to assess interobserver repeatability. For automated CT scoring, histogram characteristics were analysed: mean lung density (MLD), mode (the most highly represented attenuation value), standard deviation, kurtosis (sharpness of the density distribution), and skewness (asymmetry of the density distribution). CT-density scores (one for each tested threshold value) were expressed as the proportion of lung showing attenuation values above the selected threshold. Spirometry: performed as recommended by the ATS/ERS guidelines, predicted values were calculated using the European Community for Steel and Coal reference values.
Jain *et al* (2007) (54)	**Chest radiography** (dextrocardia, hyperinflation, bronchial wall thickening and dilation, mottled shadows,	**Microbiology**	Chest radiography: modified Chrispin-Norman score (no need for lateral film). 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.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
	consolidation or collapse), **HRCT**		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).
Kennedy *et al* (2007b) (55)	**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 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)
Li *et al* (2005) (56)	**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. Three technically

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			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.
Maglione *et al* (2012) (57)	**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. 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. 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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			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 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. 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* (2017) (58)	**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	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*, later modified by Puderbach *et al*. Maximum achievable total score was 25, indicating the most severe lung changes. For the purpose of quantifying the

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		antibiotics, hospital admissions), **microbiology** (sputum)	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.
Magnin *et al* (2012) (59)	**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*. 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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			scoring systems were used and slightly modified to obtain a score easy to use in routine practice. 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 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 ranged from 0 to 42 points without the points assessed for surgery.
Montella *et al* (2009a) (60)	**HRCT, MRI, body plethysmograghy**	**Microbiology**	MRI and HRCT scores: modified version of Helbich *et al*. The severity of mosaic perfusion was excluded as it could not be assessed by morphological MRI. The

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	(FEV_1 and FVC % predicted)		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" considered the number of missing segments. Six lobes were examined; the lingula was scored separately. 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. Body plethysmography: performed according to ATS criteria. FEV_1 > 85% predicted was considered normal.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Montella *et al* (2009b) (61)	**HRCT, MRI**	**None**	Same as above for HRCT and MRI.
Santamaria *et al* (2008) (62)	**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. 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 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. Spirometry: FEV_1 of > 85% predicted was considered normal. Microbiology: deep throat or sputum cultures were obtained.
Tadd *et al* (2019) (63)	**CT** (Brody and Bhalla scores)	None	CT: Patients were assessed for the presence and extent bronchiectasis, bronchial wall thickening, atelectasis, mucous plugging, and air trapping, using the Brody and Bhalla scoring systems. If present, each abnormality was designated as mild-moderate if the extent was <50% of the lobe, and moderate-severe ≥50% of the

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			lobe. The relative frequencies and lobar distributions of the changes were described. CT changes were annotated for all five lobes of the lung, with the lingula classified as an additional sixth lobe. Bronchiectasis was identified when the outer edge bronchus-artery cross-sectional area ratio was greater than 1, or the bronchus was non-tapering as it approached the pleura, assessed subjectively. Bronchial wall thickening was identified when airway walls were thicker than healthy airways, assessed subjectively. Mucous plugging was identified when there was a high-density occlusion seen in an airway, or tree-in-bud appearance in small airways. Trapped air was identified on expiratory images only as an area of reduced signal intensity compared to healthy lung.
Main study outcome: Microbiology			
Alanin *et al* (2015) (64)	**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'. 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;

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
			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.
Cohen-Cymberknoh *et al* (2017) (65)	**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, 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). 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
Roden *et al* (2019) (66)	**Microbiology** (mean daily alteration, yearly rate)**, spirometry** (FEV_1 and FVC % predicted)	None	Microbiology: Microbiological cultures were performed in line with recommendations for the work-up for CF specimens. The mean daily rate of alteration (MDRA) was calculated based on number of follow-up visits of patients (without baseline), and number of changed species (loss or gain) at each visit compared to the previous visit of patient , and time in days between each visit and previous visit. The yearly rate (MRA) describes the fluctuation and persistence of species in the individual patient. Spirometry: Parameters were expressed as FEV_1 % predicted and FVC % predicted estimated using the Global Lung Function Initiative reference equations (126).
Rogers *et al* (2013) (67)	**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.
Main study outcome: Anthropometry			
Goutaki *et al* (2017) (68)	**Anthropometry** (height, BMI z-scores)**, spirometry** (FEV_1, FVC z-scores)	None	Anthropometry: age- and sex-adjusted height and BMI z-scores, based on international reference values from the WHO and national reference values. For patients aged <20 years, height and BMI z-scores were calculated based on the exact age-specific references. For patients aged ≥20 years, height z-scores were calculated based on the reference values for 19-year-olds; these describe final adult

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			height. BMI z-scores were also calculated for adults, based on the reference values for 19-year-olds, because no BMI z-score references presently exist for adults. Short stature was defined as a height z-score ≤-2; underweight, as a BMI z-score ≤-2; and overweight, as a BMI z-score ≥2, according to the definitions used by WHO. Spirometry: GLI reference values were used to calculate age, sex, ethnicity, and height-adjusted z-scores for FEV_1 and FVC values. All lung function measurements were checked for quality, and since 2005, they have been performed according to ERS/ATS guidelines.
Svobodova *et al* (2013) (69)	**Anthropometry** (height SD, weight, BMI)	None	Anthropometry: data were converted into a standard deviation score (SDS) of body height, according to the latest available normative data of the background population
Main study outcome: Health-related quality of life			
Alanin *et al* (2017) (70)	**HRQoL** (SNOT-22 score)**, microbiology** (lung infection status, bronchoalveolar lavage culture, sputum culture, precipitins against Pseudomonas)**,**	None	HRQoL: SNOT-22 contains 22 questions which evaluate the effect of CRS on HRQoL. The maximum score is 110. The questionnaire has been validated to evaluate the outcome after ESS in CRS patients. Microbiology: Bronchoalveolar lavage was performed in conjunction with ESS as described. Adjuvant therapy included 2 weeks of systemic antibiotic therapy according to susceptibility testing of the bacteria cultured from the bronchoalveolar lavage and/or sinuses, 2x daily nasal irrigations with saline, and topical nasal steroids for at least 3 months. Lung bacteriology was based on sputum samples or

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	spirometry (FEV_1 and FVC % predicted), **anthropometry** (BMI)		bronchoalveolar lavage fluid. Lung infection status was assessed by modified CF Leeds criteria. The % bacteriologically positive lung samples with the dominant pathogen 12 months before surgery was compared to % positive samples during follow-up. Normal values of precipitins are 0 or 1, while 2 precipitins are considered abnormal. Chronic infection was defined as abnormal precipitins with a positive sample for Pseudomonas from the lower airways. Spirometry: ATS standards.
Behan *et al* (2017) (71)	**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. 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,

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			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 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.
Carotenuto *et al* (2013) (72)	**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

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
			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.
Ioannou *et al* (2020) (73)	**HRQoL** (QOL-PCD questionnaire, SF-36)	**Spirometry** (FEV_1, FVC z-scores)	HRQoL: The Greek version of the adult QOL-PCD questionnaire included 40 questions that compose 10 sub-scales: physical functioning (n = 5), vitality (n = 3), emotional functioning (n = 5), health perception (n = 4), treatment burden (n = 4), upper respiratory symptoms (n = 4), lower respiratory symptoms (n = 6), role (n = 4) social functioning (n = 3), hearing symptoms (n = 2). Higher scores in each subscale represent increased HRQoL.
Kenis Coskun *et al* (2019) (74)	**HRQoL** (PCD-QOL, Zerit caregiver burden scale)	**Spirometry** (FEV_1, FVC, PEF % predicted)**, anthropometry** (BMI z-score), **microbiology** (presence of	Zerit caregiver burden scale: has been widely used in investigating the caregiver burden of various chronic childhood diseases and genetic conditions. It contains 22 questions which are scored with a 5-point Likert scale. Higher scores indicate a higher burden, and the maximum score is 88.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		Pseudomonas aeruginosa)	
Maglione *et al* (2014b) (75)	**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.
McManus *et al* (2003) (76)	**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 pain and general health, which are primarily physical, and energy/vitality, social functioning, role emotional and mental health, which are primarily mental. The 8

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			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 every day' (scored 4).
McManus *et al* (2006) (77)	**HR-QoL** (SGRQ: symptoms, activity, impacts; SF-36 questionnaire: PCS, MCS; General Health Questionnaire; 'Big Five' personality dimensions, stigma questionnaire)	**None**	Same as above for SGRQ and SF-36. 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. Stigma questionnaire: study-specific measure. Used the stigma sub-scale of the 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.
Pifferi *et al* (2010) (78)	**HRQoL** (SGRQ and SF-36)	**Treatment** (daily physiotherapy, regular antibiotics, regular	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,

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		bronchodilators, intermittent bronchodilators, mucolytics, surgical procedures)	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 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

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
			diagnosis were scores from 1=greatly worsened to 5=greatly improved.
Valero-Moreno *et al* (2020) (79)	**HRQoL** (Psychological Well-Being Scale for Adolescents (BIEPS-J), Rosenberg Self-Esteem Scale (RSE), Hospital Anxiety and Depression Scale (HADS))	**Spirometry** (FVC, FEV_1, FEV_1/FVC % predicted)	HRQoL: Psychological Well-Being Scale for Adolescents (BIEPS-J) measures psychological well-being on 4 subscales (situation control, psychosocial bonds, self-acceptance and projects). It consists of 13 items, with 3 answer options: "agree", "neither agree nor disagree" and "disagree". It has an overall emotional well-being score, which is the total of all the scores. Rosenberg Self-Esteem Scale (RSE) focused on feelings of respect for and acceptance of oneself. It consists of 10 items (a Likert format, ranging from 1 -Strongly disagree, to 4—Strongly agree), focused on feelings of respect for and acceptance of oneself. The total score ranges from 10 to 40 points, distinguishing between low (scores less than or equal to 29) and high (equal to or greater than 30) self-esteem. Hospital Anxiety and Depression Scale (HADS) evaluate cognitive clinical anxiety and depression, as opposed to the somatic clinical profile. It is divided into two dimensions: the anxiety subscale (HADS-A) and the depression subscale (HADS-D). Adding the scales of anxiety and depression provides an overall score for emotional distress. Scores between 0–6 represent 'no anxiety', 7–9 'anxiety possible', over 10 'anxiety probable'. In depression, 0–5.4 represent no depression, 5.5–7.5 depression possible, over 7.5 depression probable and for emotional distress, below 15.5 no emotional distress, and over 15.5 emotional distress probable.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Whalley *et al* (2006) (80)	**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.
Zengin Akkus *et al* (2019) (81)	**HRQoL** (Ages and Stages Questionnaire for Turkish children (ASQ-TR), Ages and Stages Questionnaire: Social-Emotional (ASQ:SE), Child Behavior Checklist for ages 1.5 to 5 years (CBCL/1.5–5)), **sleep**	None	HRQoL: Ages and Stages Questionnaire (ASQ) was administered via parent interviews in conjunction with the literature. ASQ has 19 age-specific sub-questionnaires assessing the development of children in terms of communication, gross motor skills, fine motor skills, problem solving, and personal-social skills. Ages and Stages Questionnaire : Social-Emotional (ASQ:SE) is a screening tool designed to be completed by parents to assess their children's social-emotional behaviours in terms of self-regulation, compliance, communication, adaptive behaviours, autonomy, affect, and interactions with people. Child Behavior Checklist for ages 1.5 to 5 years (CBCL/1.5–5), which is the extended form of the checklist for the children between the ages 2 and 3, is designed to be completed by parents to score their own child's behaviours. CBCL/1.5–5 has seven syndrome scores: (i)

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	(Pediatric Sleep Questionnaire (PSQ))		emotionally reactive, (ii) anxious/depressed, (iii) somatic complaints, (iv) withdrawn, (v) sleep problems, (vi) attention problems, and (vii) aggressive. The combination of emotionally reactive, anxious/depressed, somatic complaints, and withdrawn scores constitute the "internalising problems score" and the combination of attention problems and aggressive scores constitute the "externalising problems score". Sleep: Pediatric Sleep Questionnaire (PSQ) is a tool to evaluate sleep-related breathing disorders in children. PSQ is composed of 22 items evaluating frequency and severity of snoring, apnoea at night sleep, breathing difficulty during sleep, daytime sleepiness, attention deficit, hyperactivity, and other paediatric obstructive sleep apnoea symptoms. Parents of children with PCD completed the validated Turkish version of Pediatric Sleep Questionnaire for the assessment of sleep related breathing disorders.
Main study outcome: Sleep disorder			
Cohen-Cymberknoh *et al* (2019) (82)	**Sleep questionnaires** (Sleep disturbance scale for children (SDSC), Pittsburg Sleep Quality Index (PSQI), Epsworth	**Spirometry** (FEV_1 % predicted)	PedsQL yields information on the physical, emotional, social and school functioning of the child during the previous 4 weeks. Abnormal scores are defined as those lower than the standard error of measurement. SDSC instrument categorises sleep disorders in children over the past 6 months in 6 subdomains to the score (disorders of initiating and maintaining sleep, sleep breathing disorders, disorders of arousal, sleep-wake transition disorders, disorders of excessive somnolence, and

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	Sleepiness Scale (ESS)), **HRQoL** (Pediatric Quality of Like Inventory (PedsQL), QOL-B)		sleep hyperhidrosis). The average global score in the general paediatric population is 35. QOL-B contains several different scales, including symptoms, physical, social and emotional functioning. PSQI questionnaire assesses sleep quality and disturbances over a 1-month time interval that has been previously used in CF patients. There are 7 component scores: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. A value of >5 is regarded as evidence of poor sleep quality. ESS respondents are asked to rate their usual chances of dozing off or falling asleep while engaged in eight different activities. Values of >10 are considered as an indication of excessive daytime sleepiness, and values between 5 and 10 indicate increased normal range daytime sleepiness. In children, parents completed the ChildHood Adenotonsillectomy Trial modified ESS.
Oktem *et al* (2013) (83)	**Body plethysmography** (FVC, FEV_1 and FEV_1/FVC % predicted), **sleep questionnaire**, **PSQI** (score, poor sleepers,	**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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	good sleepers), **polysomnography, HRCT**		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 child's sleep, restless sweating, blue colour during sleep, parental shaking for apnoea).
Santamaria *et al* (2014) (84)	**Respiratory polysomnography** (obstructive apnoea index, central apnoea index, hypopnoea index, apnoea–hypopnoea index, oxygen desaturation	**Spirometry** (FEV_1, FVC and FEF_{25-75} % predicted), **anthropometry** (BMI), **treatment** (n antibiotic courses in the last year), **microbiology** (positive	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.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	index, mean oxygen desaturation %, mean and nadir oxygen saturation %), **sleep questionnaire** (Sleep Disturbances Scale for Children), **HRCT**	sputum cultures in the last year)	The total score ranges between 26 and 130, and higher scores indicate more severe disturbances. HCRT: modified Helbich score.
Sismanlar *et al* (2018) (85)	**Sleep** (Pediatric Sleep Questionnaire, home sleep testing)**, attention deficit** (Stroop test, Conner's parents and teachers rating score)	**Spirometry** (FEV_1, FVC, FEV_1/FVC, FEF_{25-75} % predicted), **radiography** (presence of bronchiectasis, peribronchial wall thickening, atelectasis)	Sleep: Turkish validated Pediatric Sleep Questionnaire (PSQ) was completed by the parents for assessing sleep habits and quality (187). Home sleep testing (HST) is a simple, portable and easy accessible test for evaluating sleep, and it could be used safely in children. Attention deficit: Stroop test was performed, which is commonly used for evaluating selective attention, cognitive flexibility, and inhibitory control (190). Turkish validated Conner's parents (CPRS) and teacher (CTRS) rating score were used. In CPRS, there were 48 questions for evaluating children's attitude and behaviour at home. In CTRS, there were 28 questions for children's assessment of behaviours in the school. There were subscales for: inattention, hyperactivity, oppositional defiant disorder, and conduct disorder according to scales. Performance was assessed in five stages as the time (in seconds). Scoring was

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			based on how the child completed each reading as well as reading time, correction or errors made. Turkish validated Conner's parents (CPRS) and teacher (CTRS) rating score were used. In CPRS, there were 48 questions for evaluating children's attitude and behaviour at home. In CTRS, there were 28 questions for children's assessment of behaviours in the school. There were subscales for: inattention, hyperactivity, oppositional defiant disorder, and conduct disorder according to scales.
Main study outcome: Inflammatory markers			
Bush *et al* (2006) (86)	**Inflammatory markers** (IL-8 concentration), **sputum biophysical and transport properties** (dynamic viscoelasticity, wettability, cohesivity, interfacial tension, solids composition, DNA, IL-8 concentration, cough	**Spirometry** (FEV_1 and FVC % predicted), **microbiology** (chronic infection with *Pseudomonas aeruginosa*)	Spirometry: performed according to ATS guidelines. 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 non-shearing 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.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	transportability)		
Cockx *et al* (2017 a) (87)	**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.
Cockx *et al* (2017 b) (88)	**Inflammatory markers** (monocytes, CCR1, CCR2, CCR5, BLT1 and FPR1, CL2, fMLP, C5a, LTB4, CD14, CD16, IL-1β, TNF-α, CCL3, CCL5, CCL18 and CCL22)	**Spirometry** (FEV_1 and FVC % predicted), **microbiology** (sputum)	Inflammatory markers: The induced cytokines and chemokines present in the supernatants after 24 h of stimulation were measured by ELISA. Analysed non-classic monocytes by flow cytometry to determine whether a shift between those monocyte subgroups can be observed in PCD patients. Phagocytic capacity of monocytes was tested with fluorescent beads coated with S. aureus.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
Paff *et al* (2017) (89)	**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,	**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. LRTI-VAS: Symptoms were measured using a modified lower respiratory tract infection visual analogue scale (LRTI-VAS). Four of five symptom domains were

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	monocytes), **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**		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.
Ratjen *et al* (2016) (90)	**Inflammatory markers from sputum** (IL-8, neutrophil elastase activity, total cell count, % neutrophils,	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. Inflammatory markers and microbiology: obtained from spontaneously expectorated sputum.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	absolute neutrophils, bacterial density), **spirometry** (FEV_1, FVC and $FEF_{25–75}$ % predicted, change in FEV_1 and FVC from baseline in %, pulmonary exacerbation score), **microbiology** (presence of the pathogens in sputum)		
Zihlif *et al* (2006) (91)	**Inflammatory markers from exhaled breath condensate and sputum** (IL-8, LTB4 and 8-isoprostane, sputum neutrophil	**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.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	count)		
Main study outcome: Exacerbations			
Kobbernagel *et al* (2020) (92)	**Number of exacerbations, spirometry** (FEV_1, FVC, FEF_{25-75} % predicted), **body plethysmography** (RV, RV/total lung capacity, airway residence % predicted), **MBW** (LCI, $S_{cond}*V_T$, $S_{acin}*V_T$), **HRQoL** (QOL-PCD), **inflammatory markers** (white blood cells, C-reactive protein, interleukin	**Pulse oximetry saturation** (%), **respiratory rates** (breaths per minute), **anthropometry** (BMI)	Respiratory exacerbation was defined as any respiratory tract symptoms leading to initiation of systemic antibiotics, irrespective of the results of bacterial culture, or decline in percent of predicted FEV_1 of ≥10% points relative to the average of %predicted FEV_1 at screening and randomisation, whether antibiotics were prescribed or not. HRQoL: Three domains of the QOL-PCD questionnaire were measured: respiratory symptoms, sinus symptoms, and ear and hearing symptoms.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	1β, 8 and 10, granulocyte-colony stimulating factor, tumour necrosis factor α, growth-regulated oncogene α, monocyte chemoattractant protein-1), **microbiology**		
Piatti *et al* (2020) (93)	**Number of exacerbations, CT** (modified Bhalla score, % bronchiectasis, severity of bronchiectasis, BSI, FACED, eFACED)**,** **spirometry** (FEV_1, FVC % predicted)**,**	**Anthropometry** (BMI)	Exacerbation: defined as indicated by expert consensus. Since the median of exacerbations was 2 per year prior to the analysis in the study, patients were divided into two groups: Low-EXAC < 2/year and High-EXAC ≥2/year. CT: scores were classified according to modified Bhalla scoring system, BSI, FACED and e-FACED scores. Lingula was considered as separate lobe. If lobectomy had been performed a severity score of 3 was assigned to the missing lobe by arbitrary definition and distribution was presumed diffuse. The mean score for all lobes for each abnormality was calculated and lobar predominance was assessed. The CT scores ranged between 0 and 48. BSI identifies patients at risk of future mortality, hospital admissions and exacerbations; FACED classifies the severity of

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	microbiology (colonisation by *Pseudomonas aeruginosa*)		bronchiectasis according to 5-years prognosis; e-FACED detects patients with more frequent exacerbations. Classification of severity was stratified into mild, moderate, and severe according to the original Authors designations. Diagnosis of bronchiectasis was based on criteria by Naidich *et al*. Microbiology: Chronic bronchial infection was defined as the isolation of the same pathogen in sputum culture on 2 or more occasions, at least 3 months apart in a 1-year period. Patients were classified as non-colonized by *Pseudomonas aeruginosa* colonization if the pathogen had never been cultured or had been cultured only once, and as colonized if they showed at least 2 positive sputum cultures for *Pseudomonas* in 1 year (3 months apart). Spirometry: according to the ATS/ERS guidelines. Volumes and flows were considered as normal when >80% of the expected value. The most recent spirometry was considered.
Main study outcome: Exercise testing			
Loomba *et al* (2017) (94)	**Spirometry** (FEV_1, FVC, FEV_1/FVC and FEF_{25-75} % predicted), **exercise testing** (peak VO_2 absolute values and %	**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. Ventilatory data were obtained every 15 seconds.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	predicted, peak $EtCO_2$, exercise time, resting O_2 saturation, % increase in blood pressure, arrythmia during exercise test)		
Madsen *et al* (2013) (95)	**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	**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 was 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 (115). For children, the reference equations of Koopman *et al* were used for DLco and Zapetal *et al* for whole-body plethysmography, except sRaw for which the reference equation of Kirby *et al* was used. For adults (>18 years), reference

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	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-QoL** (study-specific questionnaire)		equation of Cotes *et al* and Quanjer *et al* 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 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.
Ring *et al* (2018)	**Exercise testing**	**Anthropometry**	Exercise peak: Testing was performed using an ergometer bike with step

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
(96)	(Peak oxygen uptake (VO_{2peak}) in mL/kg/min, z-score and %abnormal, single-breath diffusing capacity for carbon monoxide (DL_{CO}))**, spirometry** (FEV_1, FVC and FEF_{25-75} z-scores)	(height, BMI z-score), **microbiology**	increments determined according to the modified Godfrey protocol. A national reference material of VO_{2peak} data from 937 healthy Danish children and young adults was used. DL_{CO} test was performed as a safety precaution and to exclude an obvious oxygen uptake limitation before the exercise test. The reference equation by Koopman *et al* was applied. Spirometry: "all-ages" reference equations were used for FVC, FEV_1, and $FEF_{25–75}$. All pulmonary function tests were performed according to ATS and ERS recommendations.
Simsek *et al* (2018) (97)	**Spirometry** (FEV_1, FVC, FEV1/FVC, FEF_{25-75} % predicted), **exercise testing** (aerobic performance (modified shuttle walk test, resting heart rate, resting SpO_2%), anaerobic performance (muscle	**Anthropometry** (weight, height and BMI z-scores)	Spirometry: performed in sitting position, and the best of at least three technically acceptable manoeuvres were recorded. An FEV1 >85% predicted was considered normal. Physical activity level: determined using Bouchard 3-Day Physical Activity record. In the activity record, a day was divided into 15-min intervals, and energy expenditure was qualified on a scale from 1 to 9. Approximate median energy cost for each of the 9 categories in kcal/kg/15 min was used to compute the daily energy expenditure for each. The mean value from 3 days was considered for the analysis. Anaerobic performance: muscle power sprint test (MPST) was used, with subjects performing 15-m sprints 6 times at maximum pace with 10 seconds of recovery.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	power sprint test, hand grip strength, quadriceps muscle strength, mean anaerobic power)), **physical activity level** (mean kcal per day)		Hand grip strength (HGS) and quadriceps muscle strength (QMS) in sitting while elbow in flexion and QMS was evaluated in sitting while knee in extension. Each muscle group was tested bilaterally, and each muscle's test was repeated for three times. Average value of three reproducible attempts was recorded in Newton. The mean value of right and left sides was calculated. Both were presented as % predicted values. Aerobic performance: 15-level modified shuttle walk test (MSWT) was considered completed when subjects were unable to maintain the required speed, fail to achieve a shuttle in the time allowed, to have a SpO2 of < 75%, and to attain maximal heart rate. The distance completed was recorded in meters.
Valerio *et al* (2012) (98)	**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** (total time spent in physical activity,	**Anthropometry** (BMI and BMI SDS)	Spirometry: according to standard spirometric techniques. 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 (> 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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	vigorous physical activity)		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.
Wells *et al* (2011) (99)	**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), a**nthropometry** (height, mass, lean body mass), **Habitual Activity Estimation Scale questionnaire**	Spirometry: according to standard spirometric techniques and expressed as % predicted value for height and gender. Habitual Activity Estimation Scale questionnaire: was used as an estimation of activity levels as previously described and validated in this population. 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
Main study outcome: Others			
Joensen *et al* (2014) (100)	**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. Exhaled breath sampling: 2 measurements per patient were performed with an interval of 5 minutes between them.
Kawakami *et al* (1996) (101)	**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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
			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) (102)	**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 lythoptysis: 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.
Marino *et al* (2019) (103)	**Nutrition** (vitamin D, selenium, zinc, copper, ferritin, folate, vitamin B12, vitamin B6, iron, transferrin, transferrin iron saturation, haemoglobin, albumin, calcium, phosphate, magnesium, low	None	Spirometry: performed according to ERS/ATS guidance. GLI equations were used to estimate z-scores for FEV_1 and FVC; ethnicity specific equations were used where available. Anthropometry: performed and recorded in accordance with WHO guidelines. Moderate malnutrition was defined as a height-for-age, weight for height, BMI or FFMI of ≤ -2 z-scores below the mean of the WHO child growth standards. Reference nutrient intake (RNI) for protein and estimated average requirements (EARs) for energy. As recommended by the Scientific Advisory Committee on Nutrition in the United Kingdom (SACN), insufficient protein was defined as an intake <100% of the lower reference nutrient intake (LRNI—meeting nutrient requirements for 2.5% of population), sufficient intake was between the LRNI 100% and ≤200% of the RNI and excessive intake ≥200% of the RNI.

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	energy intake), **spirometry** (FEV_1 and FVC % predicted and z-scores)**,** **anthropometry** (weight, height and BMI z-scores, fat free mass index, bioelectrical impedance spectroscopy)**,** **inflammatory markers** (c-reactive protein, alkaline phosphatase, pro-inflammatory cytokines (IL-1B, IL-2, IL-6, IL-8 & TNF-α))		
Mirra *et al* (2015) (104)	**Vitamin D** (total 25(OH)D), **body**	**Anthropometry** (BMI), **HRCT** (bronchiectasis),	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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	plethysmography (FVC, FEV_1, FEF_{25-75}, FRC, RV and FEV_1/FVC % predicted), **HR-QoL** (SGRQ), **physical activity assessment** (questionnaire), **microbiology**	**treatment** (number of courses of antibiotics)	ng/ml (<50 nmol/L) Self-reported physical activity: assessed using a previously published questionnaire by Madsen *et al*. Microbiology: chronic bacterial colonization was defined as persistence of specific bacteria for at least 6 months, with at least 3 positive cultures.
Montuschi *et al* (2014) (105)	**Breath profiles** (ethanol, methanol, saturated fatty acids, formate, lactate, acetate, leucine/isoleucine, isobutyrate, glutamine/glutamic acide)	**Spirometry** (FEV_1 and FVC % predicted), **microbiology** (sputum culture), **anthropometry** (BMI), **treatment** (inhaled medication)	Not reported – correspondence, therefore limited information available.
Noone *et al* (1999) (106)	**Clearance during cough** (mean	**Spirometry** (FEV_1 % predicted), **cough**	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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
	clearance rates (%/min)), **sputum production rate** (sputum rheology and ion content (Avg Log G, cough-clearance index, mucociliary-clearance index, Na+ content, Cl- content))	**questionnaire** (cough severity and type, amount, ease of expectoration, and nature of sputum, chest tightness, and wheezing)	(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.
Paff *et al* (2013) (107)	**Exhaled breath profile** (volatile organic compounds)	**Spirometry** (best FEV_1 and FVC in past year % predicted), **microbiology** (positive bacterial cultures by pathogens), **pulmonary exacerbations**	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

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		(number of episodes)	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.
Pifferi *et al* (2017) (108)	**Spirometry** (FEV_1, FVC, FEF_{25-75} and FEV_1/FVC z-scores)**,** **HRCT** (modified Bhalla)**, body plethysmography** (Raw, sRaw, sReff, FRC, RV, TLC and RV/TLC z-scores) **microbiology, extracellular matrix** (metalloproteinase-8 and -9, metalloproteinase tissue inhibitors)	None	HRCT: The modified Bhalla score 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. Severity class for total lung impairment (from 1 to 3) was calculated: class of severity 1 for total score of 0-6, class 2 for total score of 7-12, class 3 for total score of 13-18. Spirometry and body plethysmography: performed according to ATS guidelines. At least three reproducible manoeuvres were obtained for each patient. To be accepted, single inspiratory manoeuvres needed to have yielded virtually superimposable XY plots, and values of FRCpleth had to be within 5% of each other.
Shoemark *et al*	**FENO** (FENO50,	**Anthropometry**	Fraction of exhaled nitric oxide (FENO): J'awNO is total NO flux in the airways and

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
(2009) (109)	FENO100, FENO200, J'awNO, CalvNO)	(height, weight), **spirometry** (FEV_1 raw), **treatment** (requirement for antibiotics, inhaled corticosteroids), **microbiology** (pathogens in sputum samples), **nasal NO** (ppb)	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.
Smit *et al* (1996) (110)	**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.
Zihlif *et al* (2005) (111)	**Cough frequency** (n cough episodes), **cough symptom score**	**Spirometry** (FEV_1 % predicted), **eNO, inflammatory markers** (sputum	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

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
		neutrophil count), **microbiology** (presence of pathogens)	were expressed as percent of predicted normal values. 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. Cough symptom score: questionnaires handed to parents, with scores ranging from 0 = no cough to 5 = distressing cough
No main study outcome			
Abitbul *et al* (2016) (112)	None	**CT** (bronchiectasis), **fertility, microbiology** (sputum cultures), **spirometry** (FEV_1 % predicted)	Not described

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Boon *et al* (2014) (113)	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. Spirometry: z-scores reported according to Quanjer equations.
Eden *et al* (2019) (114)	None	**Spirometry** (FEV_1, FVC, FEV_1/FVC % predicted), **microbiology, number of exacerbations** (past 2 years)	Exacerbations were recorded as historical information and based on the answer to the BRR baseline question: “Has the patient experienced an exacerbation of bronchiectasis within the past 2 years?” Investigators at each centre had available the definition of an exacerbation as given by O’Donnell *et al* as a guideline for the response to the question.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Emiralioglu *et al* (2020) (115)	None	**Spirometry** (FEV_1, FVC, FEF_{25-75} % predicted and z-scores)**,** **anthropometry** (BMI z-scores)**,** **microbiology,** **lobectomy** (history), **CT** (bronchiectasis)	Anthropometry: BMI was calculated by dividing weight in kilograms by the square of height in meters. The z-score for BMI-for-age was obtained from the WHO AnthroPlus packet programme. Spirometry: performed in accordance with the American Thoracic Society standards.
Frija-Masson *et al* (2017) (116)	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**	Spirometry: performed according to the ERS/ATS guidelines. 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. 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).

Authors (year of publication)	Study outcomes	Population descriptors	Definition of outcome measures
		score (Modified Medical Research Council scale), **treatment** (number of courses of antibiotics (IV, oral, inhaled)), **fertility, lobectomy** (long-term oxygen use, lung transplant), **mortality**	
Knowles *et al* (2014) (117)	**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.
Noone *et al* (2004) (118)	**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.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Pifferi *et al* (2015) (119)	**None**	**Spirometry** (FEV_1, FVC and FEF_{25-75} % predicted, changes in FEV_1 and FEF_{25-75} % predicted after bronchodilator), **HRCT** (bronchiectasis (%), class total lung impairment, class extent of bronchiectasis, class severity of bronchiectasis), **microbiology** (infection with *P aeruginosa*)	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 TEM results that are not PCD-specific or that disappear upon cellular regrowth in culture.

Authors (year of publication)	**Study outcomes**	**Population descriptors**	**Definition of outcome measures**
Yiallouros *et al* (2015) (120)	**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 z score), **lobectomy** (location of resected lobe)	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.

HRCT: High-resolution computed tomography, CT: computed tomography, FEV_1: forced expiratory volume in one second, MBW: multiple-breath washout, SBW: single-breath washout, FRC: functional residual capacity, SD: standard deviation, FVC: forced vital capacity, HRQoL: health-related quality of life, QOL-PCD: Quality of life-primary ciliary dyskinesia, SF-36: Short-Form 36 Health Survey, SGRQ: St George Respiratory Questionnaire, SNOT-20: Sino-Nasal Outcome Test 20, COPD: chronic obstructive pulmonary disease, BMI: body mass index, FEF_{25-75}: forced expiratory flow at 25-75%, LCI: lung clearance index, ATS: American Thoracic Society, ERS: European Respiratory Society, IL: interleukin, IQ: intelligence quotient, LT; leukotriene, CXCL: chemokine ligand, TLC: total lung capacity, IV: intravenous, SPO_2: peripheral capillary oxygen saturation, VO_2: oxygen consumption measured during incremental exercise, $EtCO_2$:

end-tidal carbon dioxide, RV: residual volume, DL_{CO}: diffusing capacity of the lungs for carbon monoxide, VR: ventilatory reserve, Abx: antibiotics, DTG: double-tracer gas, VOC: volatile organic compounds, TNF: tumor necrosis factor, IFN: interferon, PEFR: peak expiratory flow rate, FRCpleth: functional residual capacity made by plethysmography, Pi: inorganic phosphate, PCr: phosphocreatine, ADP: adenosine di-phosphate, ATP: adenosine-5'-triphosphate

Supplementary E-table 3. Summary of study characteristics of cross-over randomised controlled trials included in this systematic review.

Authors (year of publication)	***n* PCD patients**	**Intervention**	**Reference group**	**Limitations**
Kobbernagel *et al* (2020) (92)	90	Azithromycin maintenance therapy	Placebo	Did not reach the estimated sample size of 125 patients
Paff *et al* (2017) (89)	22	Hypertonic saline	Isotonic saline	Small sample size Non-disease-specific outcomes Isotonic saline might have beneficial effect
Gokdemir *et al* (2014) (24)	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
Koh *et al* (1999) (27)	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

Authors (year of publication)	***n* PCD patients**	**Intervention**	**Reference group**	**Limitations**
Noone *et al* (1999) (106)	12	Aerosolised uridine-5'-triphosphate	Placebo (0.12% saline)	Small sample size All had bronchiectasis (disease severity) Unclear clinical significance as differences were only temporary