SYSTEMATIC REVIEW

Accuracy of Nasal Nitric Oxide Measurement as a Diagnostic Test for Primary Ciliary Dyskinesia
A Systematic Review and Meta-analysis

Adam J. Shapiro[1], Maureen Josephson[2], Margaret Rosenfeld[3], Ozge Yilmaz[4], Stephanie D. Davis[5], Deepika Polineni[6], Elena Guadagno[7], Margaret W. Leigh[8] and Valery Lavergne[9]

[1]Division of Pediatric Respiratory Medicine, Montreal Children's Hospital, McGill University Health Centre Research Institute, Montreal, Quebec, Canada; [2]Division of Pediatric Pulmonology, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, Pennsylvania; [3]Division of Pediatric Pulmonology, Seattle Children's Hospital and Regional Medical Center, University of Washington, Seattle, Washington; [4]Department of Pediatric Allergy and Pulmonology, Faculty of Medicine, Celal Bayar University, Manisa, Turkey; [5]Section of Pediatric Pulmonology, Allergy and Sleep Medicine, Department of Pediatrics, Riley Children's Hospital, Indiana University School of Medicine, Indianapolis, Indiana; [6]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, University of Kansas, Kansas City, Kansas; [7]Medical Library, McConnell Resource Center, McGill University Health Centre, Montreal, Quebec, Canada; [8]Division of Pediatric Pulmonology, Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and [9]Department of Medical Biology, Sacré-Coeur Hospital, University of Montreal, Montreal, Quebec, Canada

Abstract

Rationale: Primary ciliary dyskinesia (PCD) is a rare disorder causing chronic otosinopulmonary disease, generally diagnosed through evaluation of respiratory ciliary ultrastructure and/or genetic testing. Nasal nitric oxide (nNO) measurement is used as a PCD screening test because patients with PCD have low nNO levels, but its value as a diagnostic test remains unknown.

Objectives: To perform a systematic review to assess the utility of nNO measurement (index test) as a diagnostic tool compared with the reference standard of electron microscopy (EM) evaluation of ciliary defects and/or detection of biallelic mutations in PCD genes.

Data Sources: Ten databases were searched for reference sources from database inception through July 29, 2016.

Data Extraction: Study inclusion was limited to publications with rigorous nNO index testing, reference standard diagnostic testing with EM and/or genetics, and calculable diagnostic accuracy information for cooperative patients (generally >5 yr old) with high suspicion of PCD.

Synthesis: Meta-analysis provided a summary estimate for sensitivity and specificity and a hierarchical summary receiver operating characteristic curve. The Quality Assessment of Diagnostic Accuracy Studies-2 tool was used to assess study quality, and Grading of Recommendations Assessment, Development, and Evaluation was used to assess the certainty of evidence. In 12 study populations (1,344 patients comprising 514 with PCD and 830 without PCD), using a reference standard of EM alone or EM and/or genetic testing, summary sensitivity was 97.6% (92.7–99.2) and specificity was 96.0% (87.9–98.7), with a positive likelihood ratio of 24.3 (7.6–76.9), a negative likelihood ratio of 0.03 (0.01–0.08), and a diagnostic odds ratio of 956.8 (141.2–6481.5) for nNO measurements. After studies using EM alone as the reference standard were excluded, the seven studies using an extended reference standard of EM and/or genetic testing showed a summary sensitivity of nNO measurements of 96.3% (88.7–98.9) and specificity of 96.4% (85.1–99.2), with a positive likelihood ratio of 26.5 (5.9–119.1), a negative likelihood ratio of 0.04 (0.01–0.12), and a diagnostic odds ratio of 699.3 (67.4–7256.0). Certainty of the evidence was graded as moderate.

Conclusions: nNO is a sensitive and specific test for PCD in cooperative patients (generally >5 yr old) with high clinical suspicion for this disease. With a moderate level of evidence, this meta-analysis confirms that nNO testing using velum closure maneuvers has diagnostic accuracy similar to EM and/or genetic testing for PCD when cystic fibrosis is ruled out. Thus, low nNO values accompanied by an appropriate clinical phenotype could be used as a diagnostic PCD test, though EM and/or genetics will continue to provide confirmatory information.

Keywords: primary ciliary dyskinesia; Kartagener syndrome; nitric oxide

Primary ciliary dyskinesia (PCD) is a rare autosomal recessive disease resulting in impaired mucociliary clearance and chronic otosinopulmonary infections. Nasal nitric oxide (nNO) levels are low in PCD, and because nNO results are immediately available, these measurements are often used as a screening tool for PCD before proceeding to ciliary electron microscopy (EM), high-speed videomicroscopic analysis (HSVA), or genetic analysis for confirmatory diagnostic testing. These latter tests are expensive (U.S.$550–U.S.$2,200), can take months to complete, and sometimes yield nondiagnostic results. Inexperience in obtaining biopsy samples can lead to insufficient cilia for EM analysis, and inexperience in interpretation can lead to false-positive or false-negative EM results. Diagnostic HSVA testing can be challenging because (*1*) there is no standardization of ciliary waveform analysis, (*2*) multiple biopsies at separate visits or re-differentiation of ciliated cells in culture is required to ensure permanence of diagnostic ciliary waveform abnormalities (i.e., not arising from secondary insults such as viral infection) (1), and (*3*) interpretation of HSVA samples from healthy control subjects shows poor interobserver agreement (2). Finally, genetic testing currently can detect biallelic mutations in only about two-thirds of patients with PCD (3).

Previous researchers have examined the diagnostic testing accuracy of nNO in PCD, but many of them incorporated methodological flaws in their study design, which could affect diagnostic accuracy. These errors included using HSVA as a screening test for study entry (excluding all subjects with normal videomicroscopy from further PCD testing), incorporating nNO measurement into both index (the new test being evaluated) and reference (the chosen gold standard) standard testing (4, 5), or using imperfect reference standard testing by enrolling some subjects diagnosed with PCD

through HSVA analysis alone and not presenting data on permanence of ciliary waveform abnormalities based on repeat HSVA testing or after cellular regrowth (6). Authors of two previous meta-analyses examined the diagnostic testing accuracy of nNO in PCD, but these analyses included studies with methodological flaws (7, 8). These methodological errors included (*1*) not providing detailed information on tests used to diagnose patients with PCD (9–11); (*2*) inclusion of nonstandard EM diagnoses in the reference standard (isolated inner dynein arm [IDA] defects without microtubule disorganization [MTD] and without repeat verification of isolated IDA defects based on two separate biopsies) (12, 13); (*3*) inclusion of patients with cystic fibrosis (CF) as disease control subjects, in whom nNO levels commonly fall below PCD cutoff values, impacting diagnostic accuracy (14–17); and (*4*) using nonstandard technology or techniques for nNO measurement (6, 10, 16). Additionally, these meta-analyses did not routinely incorporate genetic results into their reference standard, even though commercial genetic testing is now a frontline clinical test for PCD.

The American Thoracic Society has supported creation of clinical diagnostic guidelines for PCD. As part of these guidelines, a robust systematic review and meta-analysis was performed to examine the diagnostic testing accuracy of nNO measurement for PCD, and the results are presented here. This review uses strict inclusion and exclusion criteria to define acceptable index and reference standard testing for PCD. The objective of this analysis was to assess if nNO measurement can be used as a diagnostic test for PCD (as opposed to only a screening test) in cooperative patients (generally >5 yr old) who have a high probability of having this disease based on a highly suggestive clinical phenotype (18) and in whom CF has been

ruled out. Specifically, the usefulness of this tool is evaluated as a replacement for the diagnostic reference standards of classic EM ultrastructural ciliary defect and/or biallelic causative mutations in PCD genes.

Methods

Data Sources and Searches

For the literature search, we consulted the following consulted databases: Africa-Wide Information (EBSCOhost, Ipswich, MA), AMED Allied and Complementary Medicine (Ovid; Wolters Kluwer Health, New York, NY), BIOSIS (Ovid), The Cochrane Library (John Wiley & Sons, Hoboken, NJ), Embase (Ovid), Global Health (Ovid), MEDLINE (Ovid), PubMed (National Library of Medicine, National Institutes of Health, Bethesda, MD), Scopus (Elsevier, Amsterdam, the Netherlands), and Web of Science (Thomson Reuters, New York, NY). We manually searched all references in the included articles to identify other potential literature of interest. The search was performed starting from all database inceptions until July 29, 2016 (*see* Appendix E1 in the online supplement).

Study Selection

Eligible studies. Selected studies evaluate the accuracy of nNO testing (index test) in cooperative patients (generally >5 yr old), who were deemed to have a high probability of having PCD on the basis of a compatible clinical phenotype compared with the reference standards of classic EM ultrastructural ciliary defect (outer dynein arm defect, outer dynein arm defect plus IDA defect, IDA defect with MTD, radial spoke or central apparatus defect) and/or biallelic mutations in known PCD genes. Articles were not excluded on the basis of language or date of publication.

Correspondence and requests for reprints should be addressed to Adam J. Shapiro, M.D., Division of Pediatric Respiratory Medicine, Montreal Children's Hospital, McGill University Health Centre Research Institute, 1001 Boulevard Décarie, BRC.5016, Montreal, QC, H4A 3J1 Canada. E-mail: adam.shapiro@muhc.mcgill.ca

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org
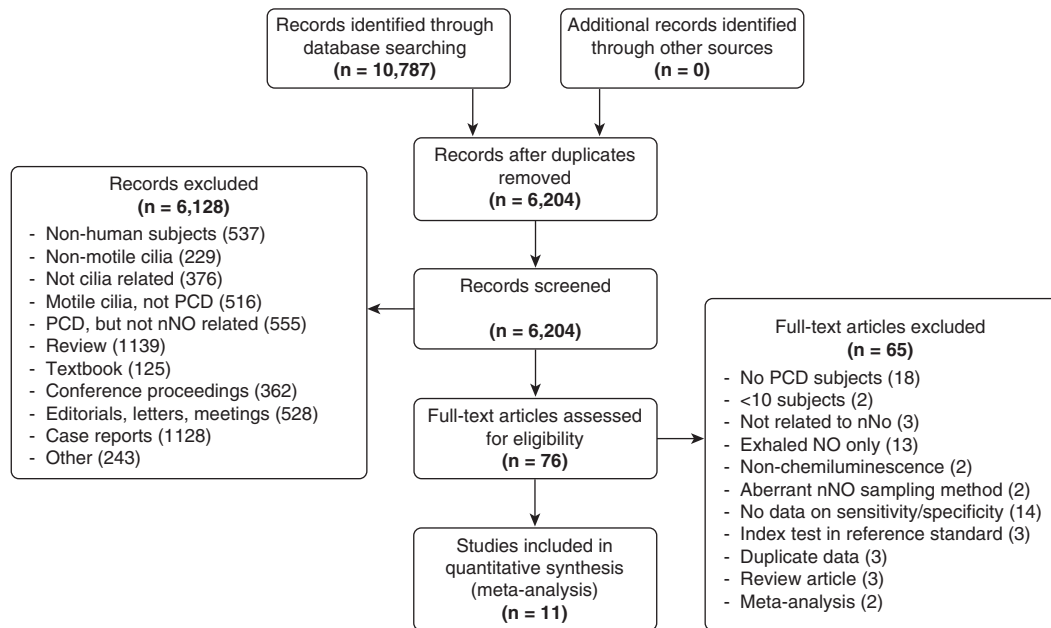
Figure 1. Summary of evidence search and selection. nNO = nasal nitric oxide; NO = nitric oxide; PCD = primary ciliary dyskinesia.

Exclusion criteria. Articles were excluded if any of the following were present:

1. Fewer than 10 patients with PCD were included in the recruited population;
2. The index test was inadequate because nNO measurement used electrochemical technology (NIOX MINO; Circassia, Uppsala, Sweden), used only nonvelum closure techniques (tidal breathing), and/or used nasal sampling flow rates outside the American Thoracic Society/European Respiratory Society recommended range (19);
3. The reference standard relied on only a single HSVA for PCD confirmation (without a second positive PCD diagnostic test or without HSVA after cellular regrowth in culture) or at least 30% of subjects had nonstandard EM defects (unrepeated, isolated IDA defects without MTD) (20);
4. Diagnostic testing accuracy was either not provided, not accurate, or not calculable; and/or
5. Index testing was incorporated into the reference standard.

Selection process. After duplicate article exclusion, two independent reviewers (A.J.S., D.P.) screened titles and abstracts to exclude nonpertinent publications. Full texts of eligible articles were assessed

for final eligibility by a team of three independent reviewers (M.J., M.R., O.Y.). Final selection was based on full-text assessment with complementary information provided by authors when needed. Three months were allowed for authors to answer e-mail queries, after which articles lacking crucial information were excluded. If the article was included but was found to contain missing information, a worst-case scenario was assumed (e.g., for unconfirmed, isolated IDA defects, patients were assumed not to have PCD). Disagreements were resolved by discussion (A.J.S., V.L.).

Data Abstraction

Two reviewers extracted data independently (A.J.S. and M.J., M.R., or O.Y.) and assessed data quality (A.J.S. and V.L.). Disagreements were resolved through discussion with a third reviewer (M.J.). nNO values derived by exhalation against resistance (ER) and breath-holding (BH) techniques were collected; BH values were accepted if ER data were unavailable. If nNO measurement techniques were unclear, authors were contacted for clarification on techniques used and the number of subjects who performed ER or BH maneuvers. All nNO measurement data are presented in nanoliters per minute. Quality assessment data were collected, including blinding to

reference or index tests, prespecification of the PCD diagnostic nNO cutoff value, and index test results as compared with the reference standard (true-positive, false-positive, true-negative, false-negative, and inconclusive result).

Quality Assessment

The Quality Assessment of Diagnostic Accuracy Studies-2 tool was used to evaluate the internal and external validity of each study (21). Risk of bias and applicability were assessed in four domains (patient selection, index test, reference standard, and flow/timing). Each item was graded as low, high, or unclear risk. The Grading of Recommendations Assessment, Development, and Evaluation system for diagnostic test accuracy (22–24) was used to analyze the certainty of evidence for each test result and for overall accuracy. In assessing the certainty of evidence, we considered the study design, risk of bias, precision, consistency, and directness.

Data Synthesis and Analysis

A bivariate model was used to calculate summary estimates for sensitivity and specificity with a generalized linear mixed model approach. Summary likelihood ratios and diagnostic odds ratios were reported. Good discrimination was defined as a positive likelihood ratio greater than 5.0 and

Table 1. Study and patient characteristics

Study Author, Year (Reference)	Location	Study Design	Total Patients* (n)	Patient Description	Prevalence of Patients with PCD (n)	Age	Male Sex (n [%])
Beydon et al., 2015 (29)	France	Cohort	86 patients suspected of having PCD	Patients included children with chronic rhinosinusitis, serous otitis media, bronchiectasis, chronic bronchitis, or situs inversus.	49 PCD total; only 44 PCD performed nNO test correctly 49 of 86 (57.0%)	PCD median = 11.4 yr (range, 7–13.9) Non-PCD median = 7.9 yr (range, 4.9–11.6)	81 of 142 (57.0%)
Boon et al., 2014 (14)	Belgium	Case–control	191 patients total 38 PCD 153 non-PCD (51 HC, 48 asthma, 54 humoral immunodeficiency)	Patients with PCD included children and adults with recurrent upper or lower respiratory tract infections with or without organ situs anomalies.	38 (NA)	Range = 5–25 yr PCD = 14.3 yr (range, 8.8–18.1) Non-PCD HC = 14.9 yr (range, 10.8–20.4), asthma = 12.1 yr (range, 9.8–16.5), humoral immunodeficiency = 10.7 yr (range, 8.2–15.6)	85 of 191 (44.5%)
Harris et al., 2014 (16)	United Kingdom	Case–control	44 patients 13 PCD 31 non-PCD (16 with symptoms, 15 HC)	Unclear	13 (NA)	Range = 6–79 yr	Not given
Leigh et al., 2013 (17) (leading site)	United States	Case–control	296 patients 149 PCD 147 non-PCD (37 asthma, 32 COPD, 78 HC)	Patients with PCD included children and adults with respiratory features suggestive of PCD (unexplained neonatal respiratory distress, year-round nasal congestion, year-round wet cough, more than five episodes of otitis media by age 2 yr, or situs anomalies, usually after cystic fibrosis and immunodeficiency excluded.	149 (NA)	PCD mean = 19.1 ± 14.8 yr Non-PCD HC mean = 20.9 ± 15.7 yr, asthma = 14.8 ± 11.5 yr, COPD = 61.1 ± 8.9 yr	139 of 296 (47.0%)
Leigh et al., 2013 (17) (other sites)	United States	Cohort	155 patients suspected of having PCD	Patients included children and adults with respiratory features suggestive of PCD (unexplained neonatal respiratory distress, year-round nasal congestion, year-round wet cough, more than five episodes of otitis media by age 2 yr, or situs anomalies, usually after cystic fibrosis and immunodeficiencies excluded.	71 of 155 (45.8%)	PCD mean = 23.3 ± 18 yr Non-PCD mean = 31.8 ± 22.3 yr	64 of 155 (41.3%)
Mateos-Corral et al., 2011 (15)	Canada	Case–control	53 patients 20 PCD 33 non-PCD (14 bronchiectasis, 19 HC)	Patients with PCD included children with sinopulmonary symptoms typical of PCD, with CF and immunodeficiency ruled out.	20 (NA)	PCD mean = 11.4 ± 3.5 yr Bronchiectasis mean = 10.9 ± 3.3 yr, HC mean = 11.0 ± 3.7 yr	26 of 53 (49.1%)
Noone et al., 2004 (30)	United States	Case–control	140 patients 69 PCD 71 non-PCD (27 HC, 44 healthy heterozygotes)	Patients with PCD included children and adults with lower airway disease with productive cough, wheeze, or shortness of breath and chronic upper airway symptoms of rhinitis/sinusitis with or without situs inversus totalis.	69 (NA)	Children with PCD median = 8 yr (range, 1–17) Adults with PCD median = 36 yr (range, 19–73) Non-PCD HC mean = 37 ± 2 yr, healthy heterozygotes = 44 ± 2 yr	PCD: 36 of 78 (46.2%)
Papon et al., 2012 (31)	France	Cohort	34 patients suspected of having PCD	Patients included children and adults with chronic upper and/or lower respiratory tract infections, bronchitis, bronchiectasis, and sinusitis.	13 of 34 (38.2%)	Mean = 32.5 yr (range, 10–72)	16 of 34 (47.1%)

(Continued)

Table 1. (*Continued*)

Study Author, Year (Reference)	Location	Study Design	Total Patients* (n)	Patient Description	Prevalence of Patients with PCD (n)	Age	Male Sex (n [%])
Piacentini et al., 2008 (32)	Italy	Case–control	35 patients 8 PCD 27 non-PCD (HC)	Patients with PCD included children with situs inversus and/or bronchiectasis and/or sinusitis.	10 PCD total; only 8 performed nNO test correctly (NA)	PCD mean = 17 yr Non-PCD = 27 school aged, mean age 7 yr	53 of 87 (60.9%)
Pifferi et al., 2011 (33)	Italy	Cohort	173 patients suspected of having PCD	Patients included children with clinical history and symptoms of PCD, without cystic fibrosis, aspiration, gastroesophageal reflux, or immunodeficiency.	48 PCD total; only 40 PCD performed nNO test correctly 48 of 173 (27.7%)	Median = 6.2 yr (range, 1 mo–17.5 yr)	105 of 209 (50.2%)
Santamaria et al., 2008 (35)	Italy	Case–control	28 patients 14 PCD 14 non-PCD (14 HC)	Unclear	14 (NA)	PCD mean = 15 yr (range, 7–27) HC mean = 16 yr (range, 7–27)	18 of 28 (64.3%)
Wodehouse et al., 2003 (34)	United Kingdom	Case–control	108 patients 42 PCD 66 non-PCD (20 bronchiectasis, 12 Young's syndrome, 18 sinusitis, 16 HC)	Unclear	42 (NA)	PCD mean = 34.2 ± 10.9 yr Non-PCD range of means = 36.2–53.2 yr	48 of 108 (44.4%)

Definition of abbreviations: CF = cystic fibrosis; COPD = chronic obstructive pulmonary disease; HC = healthy control subjects; NA = not applicable; nNO = nasal nitric oxide; PCD = primary ciliary dyskinesia.
*Number of patients included in our final analysis after excluding patients experiencing technical difficulties with nNO testing (Beydon and colleagues [29; n = 39] and Pifferi and colleagues [33; n = 3]), subjects with CF (Boon and colleagues [14; n = 50], Harris and colleagues [16; n = 6], Leigh and colleagues [17; lead site, n = 77], Mateos-Corral and colleagues [15; n = 32], Noone and colleagues [30; n = 11], and Wodehouse and colleagues [34; n = 15]), and patients with an inconclusive reference standard result (Beydon and colleagues [29; n = 56]). Additionally, uncooperative children who could perform only tidal breathing nNO measurements were excluded from the analysis (Beydon and colleagues [29; PCD, n = 5; non-PCD, n = 7], Piacentini and colleagues [32; PCD, n = 2; healthy control subjects, n = 50], and Pifferi and colleagues [33; PCD, n = 8; non-PCD, n = 28]).

a negative likelihood ratio less than 0.2 (25). A hierarchical summary receiver operating characteristic curve was constructed to describe the relationship between a continuous cutoff and accuracy. Analyses were performed using Stata version IC 14 software (StataCorp, College Station, TX) with the commands "metandi" and "metandiplot" (26). Different sources of heterogeneity other than variation in thresholds between studies were explored. A sensitivity analysis was performed to estimate the accuracy of nNO testing after excluding studies relying on EM alone as a reference standard. Other sources of heterogeneity were explored using subgroup analyses. Analyses were performed in Review Manager 5.3 (The Cochrane Collaboration, London, UK). Heterogeneity was assessed by visual inspection of the summary receiver operating characteristic curve.

General Methodology
This article follows the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols guidelines for systematic reviews and meta-analyses (Appendix E2) (27) and the *Cochrane Handbook for DTA Reviews* recommendations (28).

Results

Study Selection
In total, 10,787 records were identified through a generalized search of all publications related to PCD for use in comprehensive guidelines on PCD diagnosis. Results were not initially limited to articles investigating nNO testing. After removing duplicates, 6,204 records were screened by title and abstract, and 6,127 records not addressing nNO testing were excluded. Seventy-six full-text articles were assessed for eligibility, 65 of which were excluded (Figure 1). Twelve study populations from 11 articles were included in the quantitative synthesis (14–17, 29–35).

Study Characteristics
All 12 included studies were published between 2003 and 2015, representing the following countries: Italy (n = 3), United States (n = 3), France (n = 2), the United Kingdom (n = 2), Belgium (n = 1), and Canada (n = 1). Sample sizes ranged from 28 to 373 patients (8–149 patients with PCD, 15–153 patients without PCD). Four studies used cohort designs (prospective investigation of consecutive symptomatic referrals for PCD), and eight studies used case–control designs (retrospective comparison of previously diagnosed PCD populations with healthy and/or disease control subjects).

Population Characteristics
A total of 1,721 patients were included in these 12 studies (Table 1). In two studies, 42 patients were excluded for technical difficulties (problems with the NO analyzer, nasal obstruction, high ambient NO, or incomplete data) (29, 33). We excluded 191 patients with CF (14–17, 30, 34) to better reflect real-world practice, where CF should be ruled out before nNO testing for PCD, as well as 88 uncooperative children who could not perform nNO with velum closure techniques (32, 33). We further excluded

Table 2. Index test and reference standard characteristics

Study Authors, Year (Reference)	Index Test Characteristics*				Reference Standard Characteristics*			
	Analyzer	Flow Rate (L/Min)	Method	Cutoff (nl/Min)	PCD Diagnosis	EM Ultrastructure	Genetics	PCD Diagnosis Not Confirmed by EM and/or Genetics
Beydon et al., 2015[†] (29)	NIOX Flex, Endono 8000	0.30	Mainly ER, 5 PCD via TB were excluded	82.2	44 of 49 PCD analyzed: EM (n = 44) and/or genetics (n = 22)	ODA (n = 17) ODA+IDA (n = 5) Central pair (n = 10)	DNAI1 (n = 5) DNAI2 (n = 1) RSPH1 (n = 1) RSPH9 (n = 1) RSPH4A (n = 2) DYX1C1 (n = 2) RPGR (n = 1) Unknown total number of genes tested	Three IDA defects alone without confirmation by genetics (6.8%)
						IDA+MTD (n = 9)	CCDC39 (n = 6) CCDC40 (n = 3) Unknown total number of genes tested	
						IDA alone (n = 3)		
Boon et al., 2014[†] (14)	Eco Physics CLD88	0.30	ER	90	38 PCD analyzed: EM (n = 23) or HSVA after ciliary culture regrowth (n = 15), and/or post hoc confirmation by genetics (n = 21)	ODA (n = 19) IDA+MTD (n = 3) RSP (n = 1) Normal EM with abnormal HSVA (n = 15)	DNAH5 (n = 4) Only DNAH5 tested CCDC40 (n = 3) Only CCDC40 tested RSPH4 (n = 1) Unknown total number of genes tested DNAH11 (n = 10) Exome sequence used for 10 cases HYDIN (n = 2) CCDC65 (n = 1) Unknown total number of genes tested	Two normal EM without confirmation by genetics (5.1%)
Harris et al., 2014[†] (16)	NIOX Flex	0.30	BH	38	13 PCD analyzed: EM (n = 11) or HSVA after ciliary culture regrowth in some cases with post hoc confirmation by genetics (n = 2)	ODA (n = 5) ODA+IDA (n = 5) IDA+MTD (n = 1) Normal EM with abnormal HSVA (n = 2)	DNAH11 (n = 2) Only DNAH11 tested	0
Leigh et al., 2013[†] (17) (leading site)	Sievers 280i, Eco Physics CLD88, NIOX Flex	0.50, 0.33, 0.30	ER	76.9	149 PCD analyzed: EM (n = 143) or genetics (n = 6)	ODA (n = 87) ODA+IDA (n = 28) IDA+MTD (n = 23) CA (n = 5) Normal EM (n = 6)	DNAH11 (n = 6)	0
Leigh et al., 2013[†] (17) (other sites)	Sievers 280i, Eco Physics CLD88, NIOX Flex	0.50, 0.33, 0.30	ER	76.9	71 PCD analyzed: EM (n = 65) or genetics (n = 6)	ODA (n = 36) ODA+IDA (n = 13) IDA+MTD (n = 15) CA (n = 1) Normal EM (n = 3) Inadequate EM (n = 3)	Confirmed but not disclosed (n = 6) Unknown total number of genes tested	0
Mateos-Corral et al., 2011 (15)	Eco Physics CLD88	0.33	ER	58.5	20 PCD analyzed: EM (n = 20) with post hoc confirmation by genetics (n = 17)	ODA+IDA (n = 11) IDA+MTD (n = 4) ODA (n = 3) RSP (n = 2)	DNAH5 (n = 6) DNAH11 (n = 1) DNAI2 (n = 1) CCDC39 (n = 2) CCDC40 (n = 1) DYX1C1 (n = 3) RSHP4A (n = 1) KTU (n = 1) LRRC50 (n = 1) Two-gene panel used in one case 12-gene panel used in 12 cases 21-gene panel used in three cases 32-gene panel used in four cases	0

Table 2. (*Continued*)

Study Authors, Year (Reference)	Index Test Characteristics*				Reference Standard Characteristics*			
	Analyzer	Flow Rate (*L/Min*)	Method	Cutoff (*nl/Min*)	PCD Diagnosis	EM Ultrastructure	Genetics	PCD Diagnosis Not Confirmed by EM and/or Genetics
Noone *et al.*, 2004† (30)	Sievers 270B	0.50	BH	*100*	69 PCD analyzed: EM (n = 60) or complete clinical phenotype with *post hoc confirmation by genetics (n = 9)*	ODA (n = 31) ODA+IDA (n = 16) IDA+MTD (n = 13)	*Confirmed but not disclosed (n = 9) Only two genes tested*	0
Papon *et al.*, 2012 (31)	EVA4000	Per ATS standards	Per ATS standards	100	13 PCD analyzed: EM (n = 13)	ODA (n = 9) IDA+nexin link (n = 2) ODA+IDA (n = 1) Central pair (n = 1)		0
Piacentini *et al.*, 2008 (32)	NIOX Flex	0.30	Mainly BH, 2 PCD via TB were excluded	20.4	8 of 10 PCD analyzed: EM (n = 10)	ODA+IDA (n = 7) ODA (n = 1) IDA (n = 2)		0
Pifferi *et al.*, 2011† (33)	Eco Physics CLD88	*0.33*	*Mainly ER, 8 PCD via TB were excluded*	96	40 of 48 PCD analyzed: EM (n = 42) or HSVA after ciliary culture regrowth *with post hoc confirmation by genetics (n = 6)*	ODA+IDA (n = 23) *IDA+CA+MTD (n = 12) ODA (n = 2) IDA+MTD (n = 3) IDA (n = 2) Normal EM with abnormal HSVA (n = 6)*	*DNAH11 (n = 6) Only DNAH11 tested*	0
Santamaria *et al.*, 2008 (35)	NIOX Flex	0.28	BH	16.8	14 PCD analyzed: EM (n = 14)	ODA+IDA (n = 8) ODA (n = 1) IDA+MTD (n = 3) Central pair (n = 1) Basal body anomaly (n = 1)		One nonclassic EM anomaly without confirmation by genetics (7.1%)
Wodehouse *et al.*, 2003 (34)	LR2000	0.25	BH	50	42 PCD analyzed: EM (n = 42)	ODA (n = 21) ODA+IDA (n = 5) Transposition (n = 2) Radial spoke (n = 2) Unspecified IDA (n = 12)		12 IDA defects alone without confirmation by genetics (28.6%)

Definition of abbreviations: ATS = American Thoracic Society; BH = breath hold; CA = central apparatus defect; EM = electron microscopy; ER = exhalation against resistance; HSVA = high-speed videomicroscopic analysis; IDA+MTD = inner dynein arm and microtubule disorganization defect; ODA = outer dynein arm defect; ODA+IDA = outer dynein arm and inner dynein arm defect; PCD = primary ciliary dyskinesia; RSP = radial spoke defect; TB = tidal breathing.
*All information in italics is derived from personal communications with the studies' authors.
†Studies considered as using a combination of EM and/or genetics as the reference standard.

56 patients who had inconclusive reference standard results (29). In total, 1,344 patients were analyzed (514 patients with PCD, 830 patients without PCD). Half of the studies included mainly a pediatric population (<18–25 years old) (14, 15, 29, 32, 33, 35), and half included patients of all ages (16, 17, 30, 31, 34). The prevalence of patients with PCD in cohort studies ranged from 28 to 57% of patients included in the quantitative analyses (17, 29, 31, 33). Nine studies provided information on symptoms leading to clinical suspicion of PCD, which generally included at least one of the following: chronic rhinosinusitis, chronic otitis media, chronic bronchitis, bronchiectasis, neonatal respiratory distress, and/or organ laterality defects (mainly situs inversus totalis). CF was ruled out in six studies, and immunodeficiency

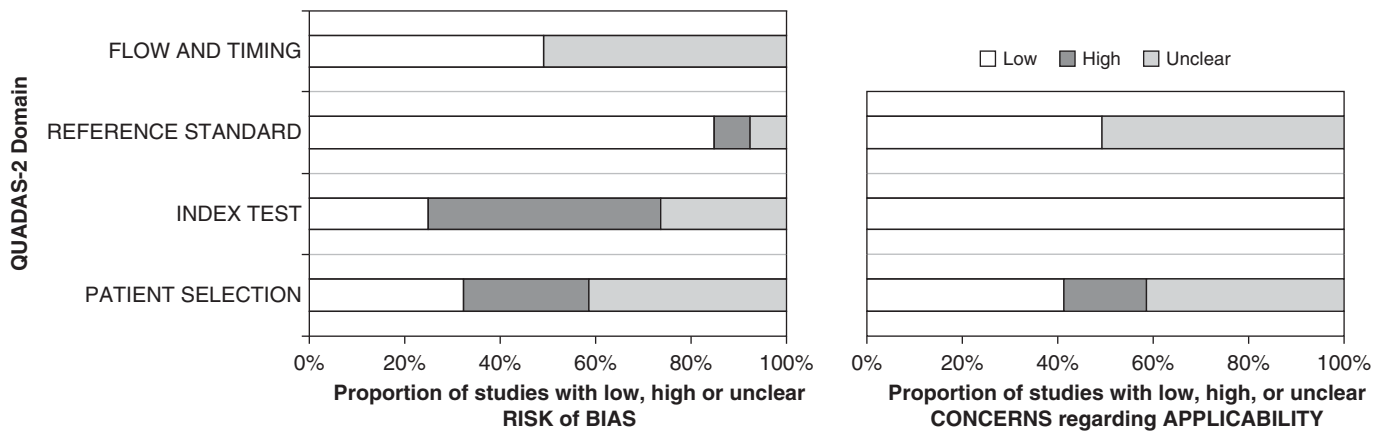was ruled out in five studies, prior to PCD testing.

Index Test Characteristics

Several different brands of chemiluminescence NO analyzers were used across the studies (NIOX Flex [Circassia], Endono 8000 [SERES Environment, Aix-en-Provence, France], CLD88 [Eco Physics, Dürnten, Switzerland], Sivers 280i [GE Analytical Instruments, Boulder, CO], EVA 4000 [Hewlett Packard Enterprise, Palo Alto, CA], LR2000 [Rochester, UK]) (Table 2). Sampling flow rates ranged from 0.25 to 0.5 L/min, but only one study included regular verification (via standard operating procedures) of sampling flow rates with direct measurement using a Gilmont flowmeter (Cole-Parmer, Vernon Hills, IL) (17). In most studies, researchers

performed device calibration according to the device manufacturer's recommendations. In six studies, researchers reported nNO measurement via ER; in five studies, researchers used BH maneuvers; and the technique used was not fully reported in one study. Diagnostic nNO cutoff values ranged from 16.8 to 100 nl/min, with a median cutoff at 76.9 nl/min.

Reference Standards Characteristics and Strategies

Electronic microscopy. All studies included ciliary EM as the sole or main reference standard (Table 2). The majority followed standard EM methodology (36). Most isolated IDA defects were either confirmed upon repeat EM study, associated with MTD on *post hoc* EM review, or confirmed

Figure 2. Assessment of validity of individual studies with Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 tool for the 12 included studies. The QUADAS-2 tool is designed to assess the quality of primary diagnostic accuracy studies and consists of four key domains evaluating the methods used in regard to patient selection, index test, reference standard, and flow of patients through the study, as well as timing of the index test and reference standard. The results presented show several studies with high risk of bias with regard to the index test domain, especially in case–control studies.

post hoc by disease-causing mutations in *CCDC39* or *CCDC40* genes. Nevertheless, Wodehouse and colleagues reported 12 patients (28.6%) as having isolated IDA defects without further specification by the authors, which increased the level of bias for this included publication (34). One basal body anomaly reported as PCD was excluded from analysis (35).

Genetic testing. Researchers in three studies reported genetic testing as part of the original reference standard (usually as a complementary tool when EM was nondiagnostic rather than as a systematic test used with all patients) (17, 29). After contacting authors, we found five additional cohorts (14–16, 30, 33) in which genetic testing was performed *post hoc* for individuals with EM defects (n = 24) or nondiagnostic EM studies (n = 32). In two

cohorts, only a single PCD gene (*DNAH11*) was tested; in one cohort, two PCD genes were tested; in one cohort, at least six PCD genes were tested; and in one cohort, 12–32 PCD genes were tested (Table 2). In the meta-analysis, patients with biallelic mutations in a PCD-causing gene, whether identified prospectively or *post hoc*, were categorized as having PCD.

Quality Assessment

Patient selection (risk of bias and applicability). Four studies had a cohort-type design (17, 29, 31, 33), whereas eight were case–control studies (Figure 2). Among the case–control studies, five used disease control subjects (14–17, 34) and three used healthy control subjects (17, 30, 32, 35). The populations examined in the

cohort studies were selected populations considered at high risk for PCD (excluding patients with CF) in whom PCD testing was being pursued.

Index test (risk of bias and applicability). In 7 of 12 studies, the nNO cutoff was not prespecified. Blinding of the index test was often not reported, but because nNO is an objective measurement, this was judged as having low impact on the risk of bias. In most studies, patients were tested when free of acute respiratory tract infection for more than 2 weeks and not around nasal instrumentation. Only nNO results from cooperative children who could perform velum closure maneuvers (via BH or ER techniques) were evaluated.

Reference standard (risk of bias and applicability). The majority of studies

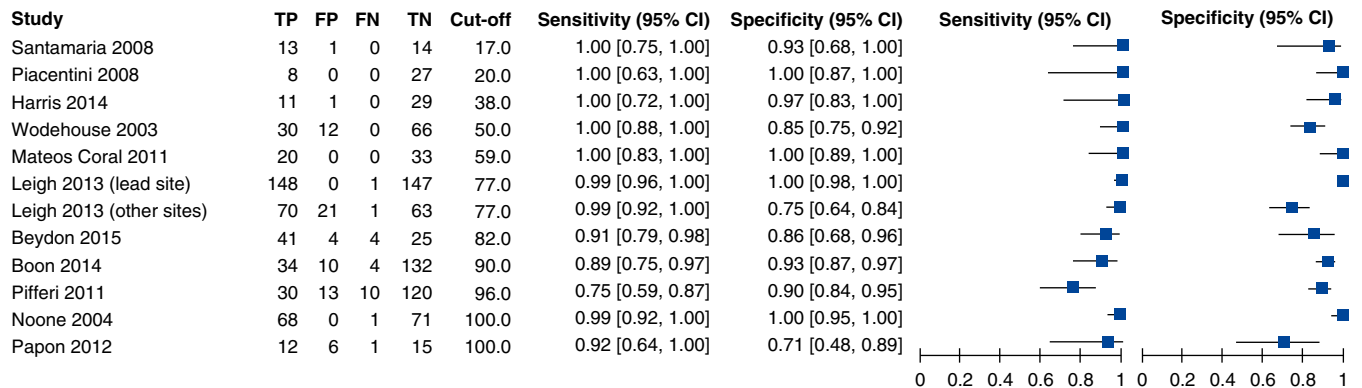| Study	TP	FP	FN	TN	Cut-off	Sensitivity (95% CI)	Specificity (95% CI)
Santamaria 2008	13	1	0	14	17.0	1.00 [0.75, 1.00]	0.93 [0.68, 1.00]
Piacentini 2008	8	0	0	27	20.0	1.00 [0.63, 1.00]	1.00 [0.87, 1.00]
Harris 2014	11	1	0	29	38.0	1.00 [0.72, 1.00]	0.97 [0.83, 1.00]
Wodehouse 2003	30	12	0	66	50.0	1.00 [0.88, 1.00]	0.85 [0.75, 0.92]
Mateos Coral 2011	20	0	0	33	59.0	1.00 [0.83, 1.00]	1.00 [0.89, 1.00]
Leigh 2013 (lead site)	148	0	1	147	77.0	0.99 [0.96, 1.00]	1.00 [0.98, 1.00]
Leigh 2013 (other sites)	70	21	1	63	77.0	0.99 [0.92, 1.00]	0.75 [0.64, 0.84]
Beydon 2015	41	4	4	25	82.0	0.91 [0.79, 0.98]	0.86 [0.68, 0.96]
Boon 2014	34	10	4	132	90.0	0.89 [0.75, 0.97]	0.93 [0.87, 0.97]
Pifferi 2011	30	13	10	120	96.0	0.75 [0.59, 0.87]	0.90 [0.84, 0.95]
Noone 2004	68	0	1	71	100.0	0.99 [0.92, 1.00]	1.00 [0.95, 1.00]
Papon 2012	12	6	1	15	100.0	0.92 [0.64, 1.00]	0.71 [0.48, 0.89]

Figure 3. Forest plot (in ascending order of nasal nitric oxide cutoff value in nanoliters per minute). CI = confidence interval; FN = false negative; FP = false positive; TN = true negative; TP = true positive.

reported that reference standards were blinded to the nNO measurements. There was no major concern regarding the technical aspects of the reference standard testing, except in one cohort study in which 39.4% patients (56 of 142) were left undiagnosed because of inconclusive reference standard results (29). The remaining 86 patients with conclusive reference standard testing were included in our meta-analysis.

Flow and timing (risk of bias). Differential verification (EM or genetic testing was performed only in patients with PCD and not in control subjects) and absence of simultaneous testing (index and reference tests were performed sequentially instead of simultaneously) were frequent, especially in case–control studies. Both of these factors may artificially increase sensitivity and specificity.

Data Synthesis

When we pooled the results of 12 studies, we observed that the bivariate analysis (average sensitivity and specificity for all thresholds) showed a summary sensitivity of 97.6% (92.7–99.2) and specificity of 96.0% (87.9–98.7), as well as a positive likelihood ratio of 24.3 (7.6–76.9), a negative likelihood ratio of 0.03 (0.01–0.08), and a diagnostic odds ratio of 956.8 (141.2–6481.5), for nNO measurements. For this analysis, isolated IDA defects were reclassified as non-PCD when feasible. Assuming a pretest probability of 35% (17, 18, 29, 31, 33), corresponding positive and negative predictive values were 92.9% (80.5–97.6) and 98.7% (95.7–99.6), respectively (Appendix E3). A forest plot presenting studies in ascending order of thresholds is presented in Figure 3. A summary hierarchical receiver operating characteristic curve illustrating how sensitivity and specificity traded off with each other as thresholds varied is presented in Figure 4.

Heterogeneity

Subgroup analysis. Sources of heterogeneity were explored using subgroup analyses. Studies presenting a lower risk of bias in different domains (e.g., using cohort-type design, disease control subjects over healthy control subjects, and prespecified nNO cutoff values) showed slightly lower diagnostic test accuracy. Interestingly, studies that systematically excluded CF prior to PCD testing (15, 17, 30, 31, 33) showed a slightly higher diagnostic accuracy than studies that did not exclude CF (14, 29, 31, 32, 34, 35) (sensitivity of
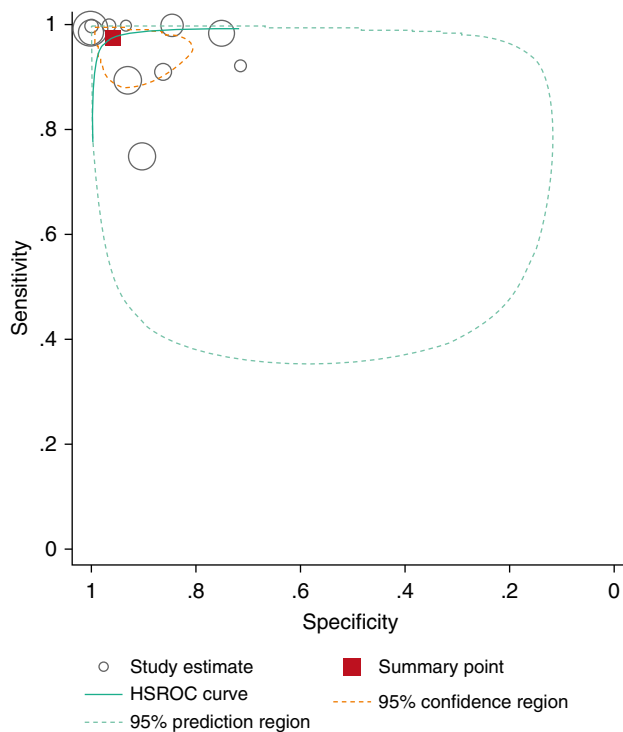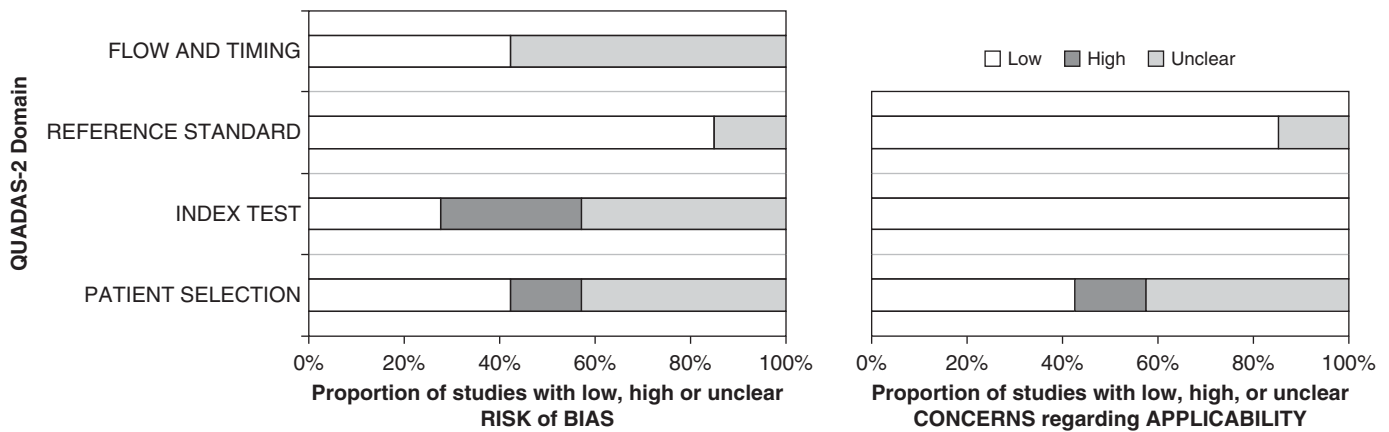

Figure 4. Hierarchical summary receiver operating characteristic curve (HSROC) for the 12 included studies.

	Risk of bias				Applicability		
	Patient selection	**Index test**	**Reference standard**	**Flow and timing**	**Patient selection**	**Index test**	**Reference standard**
Beydon 2015	Low	High	Unclear	Low	Low	Low	Low
Boon 2014	Unclear	Low	Low	Unclear	Unclear	Low	Low
Harris 2014	Unclear	High	Low	Unclear	Unclear	Low	Low
Leigh 2013 (leading site)	Unclear	Unclear	Low	Unclear	Unclear	Low	Low
Leigh 2013 (other sites)	Low	Low	Low	Low	Low	Low	Low
Noone 2004	High	Unclear	Low	Unclear	High	Low	Unclear
Pfierri 2011	Low	Unclear	Low	Low	Low	Low	Low

Figure 5. Assessment of validity of individual studies with Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 tool for the seven included studies comparing nasal nitric oxide to an extended reference standard of electron microscopy and/or genetics. The QUADAS-2 tool is designed to assess the quality of primary diagnostic accuracy studies and consists of four key domains evaluating the methods used with regard to patient selection, index test, reference standard, and flow of patients through the study, as well as timing of the index test and reference standard. The results presented show that the 7 selected studies were at lower risk of bias and concern regarding applicability than the initial 12 analyzed studies presented in Figure 2.

97.7% vs. 95.1% and specificity of 98.5% vs. 91.4%, respectively).

Sensitivity analysis. The most relevant source of heterogeneity was the strategy used for the reference standard of PCD disease (EM alone vs. extended reference standard combining EM and/or genetic testing). Thus, we performed a sensitivity analysis including only the seven studies with the extended reference standard of EM defects and/or genetic diagnoses (14–17, 29, 30, 33), which included 1,086 patients (430 patients with PCD, 656 patients without PCD). Globally, these seven studies were at lower risk of bias than the whole group (Figure 5), with proportionally more cohort-type studies, less using asymptomatic patients as their control group, and more studies prespecifying their nNO cutoff. Pooled analysis showed a summary sensitivity of 96.3% (88.7–98.9) and specificity of 96.4% (85.1–99.2), as well as a positive likelihood ratio of 26.5 (5.9–119.1), a negative likelihood ratio

of 0.04 (0.01–0.12), and a diagnostic odds ratio of 699.3 (67.4–7,255.9), when we compared nNO with the extended reference standard of EM defects and/or biallelic genetic mutations (Figure 6). According to Grading of Recommendations Assessment, Development, and Evaluation methodology, the overall certainty of evidence was moderate when we evaluated studies comparing nNO with an extended reference standard of EM and/or genetics (*see* Table 3).

Discussion

In this meta-analysis, the diagnostic testing accuracy of nNO is excellent when compared with EM, and only slightly lower when compared with the extended reference standard of EM and/or genetic testing. Both EM and genetic analysis are imperfect reference standard PCD tests, with currently estimated sensitivity at 0.70 (3, 37), and

each of these detecting PCD cases that can be missed by the other test. Additionally, these reference standard tests can frequently provide nondiagnostic results, with up to 40% of clinical biopsies showing inadequate cilia for EM analysis (38) and up to 43% of genetic testing detecting monoallelic mutations or variants of unknown significance (39). Conversely, nNO measurement is a highly feasible test in cooperative patients (generally >5 yr old), with successful measurements accomplished in more than 90% of patients in this meta-analysis. Although nNO testing has largely been considered as a PCD screening test, this analysis shows that nNO has a diagnostic potential similar to the accepted confirmatory PCD tests of EM and genetic analysis. Thus, in populations with an appropriate clinical phenotype for PCD, where CF is ruled out, nNO measurement is a comparable PCD diagnostic test, with the added benefits of being highly feasible, painless, noninvasive,
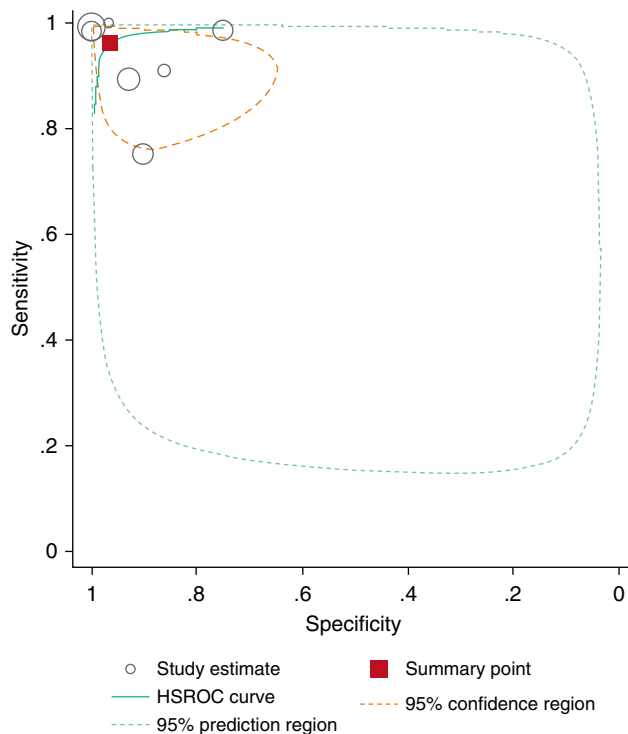
Figure 6. Hierarchical summary receiver operating characteristic curve (HSROC) for the seven studies comparing nasal nitric oxide to an extended reference standard of electron microscopy and/or genetics.

rapid, and relatively inexpensive (U.S.$25–U.S.$85) for patients. However, there are limitations to nNO testing for PCD, including high purchase cost of chemiluminescence machines, training of device operators, lack of clinical approval for nNO devices in the United States,

and the inability to rigorously test uncooperative children (generally <5 yr old).

Disease prevalence influences posttest probability, and this analysis assumes a PCD disease prevalence of 35%, as demonstrated when PCD is strongly suspected because of

the presence of a highly suggestive clinical phenotype (18). This robust phenotype is highly predictive of PCD on the basis of (*1*) unexplained neonatal respiratory distress at term birth, (*2*) year-round wet cough starting before 6 months of age, (*3*) year-round nasal congestion starting before 6 months of age, and (*4*) organ laterality defects. Although some of the studies included in this meta-analysis did not use these specific symptoms to select candidates for PCD diagnostic testing, most studies included variations of these clinical criteria. Thus, in a preselected population expressing these PCD-specific symptoms, nNO measurement is a highly accurate diagnostic test and can replace EM or genetic testing. If the prevalence is lower because of less stringent phenotype screening, the positive predictive value will be lower. For example, if the prevalence of PCD is 10% in a less-selective group, the positive predictive value of nNO testing for PCD is considerably lower at 73%. At this lower PCD prevalence, approximately one-fourth of patients with a positive nNO test will not have PCD upon confirmatory testing. Therefore, it is critical that careful selection of patients for diagnostic evaluation by nNO testing be accomplished. Otherwise, in less-enriched groups, nNO will be more useful as a triage test prior to PCD diagnostic testing, as opposed to a replacement diagnostic test. Clinicians must consider this point and appropriately screen patients for PCD-specific clinical criteria before embarking

Table 3. Summary of findings, including the seven studies comparing nasal nitric oxide with an extended reference standard of electron microscopy and/or genetics*

Outcome	Studies (n), Patients (n)	Study Design	Factors That May Decrease Quality of Evidence					Effect per 100 Patients Tested (Pretest Probability of 35%)	Test accuracy QoE	Importance
			Risk of Bias	Indirectness	Inconsistency	Imprecision	Publication Bias			
True positives (patients with PCD)	7 studies, 423 patients	Cohort and case–control-type studies	Serious[†]	Not serious	Not serious[‡]	Not serious	None	34 (31–35)	⊕⊕⊕⊖MODERATE	Critical
False negatives (patients incorrectly classified as not having PCD)								1 (0–4)		Critical
True negatives (patients without PCD)	7 studies, 636 patients	Cohort and case–control studies	Serious[†]	Not serious	Not serious[‡]	Not serious	None	63 (55–64)	⊕⊕⊕⊖MODERATE	Critical
False positives (patients incorrectly classified as having PCD)								2 (1–10)		Important
Inconclusive	7 studies, 27 patients	—	—	—	—	—	—	—	—	Important

Definition of abbreviations: PCD = primary ciliary dyskinesia; QOE = quality of evidence.
*Sensitivity, 0.96 (95% confidence interval, 0.89–0.99); specificity, 0.96 (95% confidence interval, 0.85–0.99); prevalence 35%.
[†]Four studies were case–control studies, among which one study included only healthy patients in the control group. Two studies did not prespecify the nasal nitric oxide cutoff before performing measurements and were not blinded to the reference standard.
[‡]Not downgraded for inconsistency since the residual heterogeneity was explained by the difference in the risk of bias between studies.

on PCD diagnostic investigations, including nNO testing.

Two past meta-analyses have shown findings similar to those of this analysis, but neither publication used an extended reference standard incorporating genetic testing (7, 8). Rather, included studies used varying combinations of different reference standards, including clinical phenotype, HSVA, EM, and rarely genetics. Our analysis used rigorous criteria to define reference standard testing. By contacting authors, we eliminated studies with greater than or equal to 30% isolated IDA defects and assigned greater bias to studies with 20 to 30% isolated IDA defects, because 25% of isolated IDA defects resolve on repeat EM testing (20). Through author communication, we also significantly increased reference standard data on genetic testing, which improves the generalizability of this analysis to current clinical practices in North America, where genetic testing is increasingly used in PCD diagnosis. Last, we discovered that some studies performed EM testing only if HSVA was first abnormal and often did not repeat HSVA studies on separate occasions or after cell culture. Altogether, our rigorous definition of reference standard testing greatly increases the strength of this meta-analysis.

In this analysis, we also used meticulous criteria to define the index test of nNO measurement. We restricted analysis to studies using chemiluminescence technology because only this technology is recommended for nNO measurement in PCD (40). Next, we limited our data to nNO testing only through velum closure techniques. Although tidal breathing nNO measurements are of clinical value in young children, PCD diagnostic cutoff values have not been defined for these techniques. Last, we excluded all patients with CF, who can have nNO levels below PCD cutoff values, which could affect diagnostic accuracy.

Even with our robust inclusion and exclusion criteria, this analysis has some limitations. First, despite its increasing clinical recognition, PCD is still a relatively rare disease, and our patient numbers are limited. Second, the heterogeneity of PCD reference standards poses difficulties for study generalizability. Ciliary EM alone identifies more classic cases of PCD, but it misses variant forms (41). The expense of genetic testing also creates differential verification whereby reference genetic testing is performed mainly in patients with suspected PCD and not in healthy control subjects, which affects diagnostic testing accuracy. Nonsimultaneous PCD diagnostic testing (using nNO as an initial screening test, followed by EM and/or genetic testing) may also have affected diagnostic accuracy in the selected studies, although blinding of researchers should have minimized these effects. Owing to the rapid discovery of novel PCD-causing gene mutations, most genetic panels are incomplete by the time of study publication, which further decreases the diagnostic accuracy of PCD genetic testing. However, with future discovery of novel PCD genes that result in normal ultrastructure with low nNO levels, the false-positive rate of nNO testing may decrease, and diagnostic accuracy may actually improve. Last, studies in this analysis using EM alone as the reference standard were more often designed as case–control studies, did not prespecify diagnostic nNO cutoff values, or were not blinded to nNO results during reference standard testing. Each of these factors is associated with an overestimation of diagnostic testing accuracy. Thus, although it is possible that nNO testing is actually less accurate when using an extended reference standard of EM and/or genetics, it seems more likely that studies using EM alone as the reference standard are at higher risk of bias, resulting in falsely increased diagnostic testing accuracy.

Conclusions

nNO is a sensitive and specific test for diagnosing PCD in cooperative patients (generally >5 yr old), in whom CF has been ruled out, and who have a robust clinical phenotype for PCD. The reference standard tests of EM and/or genetic analysis are imperfect tests because both lack sensitivity for PCD diagnosis. Although nNO was previously considered a PCD screening test, with a moderate level of evidence, this meta-analysis confirms that nNO testing has at least equivalent and likely better diagnostic testing accuracy than EM and/or genetic testing for PCD. Thus, we propose that nNO be considered a diagnostic test rather than a screening test in this population. Physicians must realize that normal nNO levels do not rule out PCD, and patients with highly compatible PCD clinical phenotypes but normal nNO levels should be subjected to further testing. In addition, even in individuals with a compatible clinical phenotype and low nNO, confirmatory testing with EM or genetics will yield additional diagnostic information. As more genetic causes of PCD are discovered, repeat meta-analysis will be required to evaluate the diagnostic testing accuracy of nNO measurement, and the upcoming American Thoracic Society–sponsored clinical practice guidelines on PCD diagnosis will further elaborate the accuracy of other PCD diagnostic tests. Future study of tidal breathing nNO measurement is needed to evaluate the usefulness of this noninvasive, rapid, and inexpensive test for successful PCD diagnosis in uncooperative children younger than 5 years old. ■

References

1 Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, Dell S, Eber E, Escudier E, Hirst RA, et al. European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. Eur Respir J 2017;49:1601090.

2 Kempeneers C, Seaton C, Chilvers MA. Variation of ciliary beat pattern in three different beating planes in healthy subjects. Chest 2017;151: 993–1001.

3 Zariwala MA, Knowles MR, Leigh MW. Primary ciliary dyskinesia. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Ledbetter N, Mefford HC, Smith RJH, et al., editors. GeneReviews. Seattle, WA: University of Washington, Seattle; 1993–2017 [2007 Jan 24; updated 2015 Sep 3].

4 Jackson CL, Behan L, Collins SA, Goggin PM, Adam EC, Coles JL, Evans HJ, Harris A, Lackie P, Packham S, et al. Accuracy of diagnostic testing in primary ciliary dyskinesia. Eur Respir J 2016;47: 837–848.

5 Collins SA, Behan L, Harris A, Gove K, Lucas JS. The dangers of widespread nitric oxide screening for primary ciliary dyskinesia. Thorax 2016;71:560–561.

6 Marthin JK, Nielsen KG. Hand-held tidal breathing nasal nitric oxide measurement – a promising targeted case-finding tool for the diagnosis of primary ciliary dyskinesia. PLoS One 2013;8: e57262.

7 Kouis P, Papatheodorou SI, Yiallouros PK. Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: a meta-analysis. BMC Pulm Med 2015;15:153.

8 Collins SA, Gove K, Walker W, Lucas JSA. Nasal nitric oxide screening for primary ciliary dyskinesia: systematic review and meta-analysis. *Eur Respir J* 2014;44:1589–1599.

9 Narang I, Ersu R, Wilson NM, Bush A. Nitric oxide in chronic airway inflammation in children: diagnostic use and pathophysiological significance. *Thorax* 2002;57:586–589.

10 Montella S, Alving K, Maniscalco M, Sofia M, De Stefano S, Raia V, Santamaria F. Measurement of nasal nitric oxide by hand-held and stationary devices. *Eur J Clin Invest* 2011;41:1063–1070.

11 Marthin JK, Nielsen KG. Choice of nasal nitric oxide technique as first-line test for primary ciliary dyskinesia. *Eur Respir J* 2011;37:559–565.

12 Moreno Galdó A, Vizmanos Lamotte G, Reverte Bover C, Gartner S, Cobos Barroso N, Rovira Amigo S, Liñán Cortés S, Lloreta Trull J, Busquets Monge R. Value of nasal nitric oxide in the diagnosis of primary ciliary dyskinesia [in Spanish]. *An Pediatr (Barc)* 2010;73:88–93.

13 Corbelli R, Bringolf-Isler B, Amacher A, Sasse B, Spycher M, Hammer J. Nasal nitric oxide measurements to screen children for primary ciliary dyskinesia. *Chest* 2004;126:1054–1059.

14 Boon M, Meyts I, Proesmans M, Vermeulen FL, Jorissen M, De Boeck K. Diagnostic accuracy of nitric oxide measurements to detect primary ciliary dyskinesia. *Eur J Clin Invest* 2014;44:477–485.

15 Mateos-Corral D, Coombs R, Grasemann H, Ratjen F, Dell SD. Diagnostic value of nasal nitric oxide measured with non-velum closure techniques for children with primary ciliary dyskinesia. *J Pediatr* 2011;159:420–424.

16 Harris A, Bhullar E, Gove K, Joslin R, Pelling J, Evans HJ, Walker WT, Lucas JS. Validation of a portable nitric oxide analyzer for screening in primary ciliary dyskinesias. *BMC Pulm Med* 2014;14:18.

17 Leigh MW, Hazucha MJ, Chawla KK, Baker BR, Shapiro AJ, Brown DE, Lavange LM, Horton BJ, Qaqish B, Carson JL, *et al*. Standardizing nasal nitric oxide measurement as a test for primary ciliary dyskinesia. *Ann Am Thorac Soc* 2013;10:574–581.

18 Leigh MW, Ferkol TW, Davis SD, Lee HS, Rosenfeld M, Dell SD, Sagel SD, Milla C, Olivier KN, Sullivan KM, *et al*. Clinical features and associated likelihood of primary ciliary dyskinesia in children and adolescents. *Ann Am Thorac Soc* 2016;13:1305–1313.

19 American Thoracic Society; European Respiratory Society. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am J Respir Crit Care Med* 2005;171:912–930.

20 O'Callaghan C, Rutman A, Williams GM, Hirst RA. Inner dynein arm defects causing primary ciliary dyskinesia: repeat testing required. *Eur Respir J* 2011;38:603–607.

21 Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–536.

22 Gopalakrishna G, Mustafa RA, Davenport C, Scholten RJ, Hyde C, Brozek J, Schünemann HJ, Bossuyt PM, Leeflang MM, Langendam MW. Applying Grading of Recommendations Assessment, Development and Evaluation (GRADE) to diagnostic tests was challenging but doable. *J Clin Epidemiol* 2014;67:760–768.

23 Hsu J, Brożek JL, Terracciano L, Kreis J, Compalati E, Stein AT, Fiocchi A, Schünemann HJ. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implement Sci* 2011;6:62.

24 Schünemann HJ, Oxman AD, Brożek J, Glasziou P, Jaeschke R, Vist GE, Williams JW Jr, Kunz R, Craig J, Montori VM, *et al*.; GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;336:1106–1110.

25 Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;329:168–169.

26 Takwoingi Y. Meta-analysis of test accuracy studies in Stata: a bivariate model approach. 2016 [accessed 2016 Apr 01]. Available from: http://methods.cochrane.org/sdt/

27 Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.

28 The Cochrane Collaboration. Cochrane handbook for DTA reviews. London: Author [accessed 2017 May 10]. Available from: http://methods.cochrane.org/sdt/handbook-dta-reviews

29 Beydon N, Chambellan A, Alberti C, de Blic J, Clément A, Escudier E, Le Bourgeois M. Technical and practical issues for tidal breathing measurements of nasal nitric oxide in children. *Pediatr Pulmonol* 2015;50:1374–1382.

30 Noone PG, Leigh MW, Sannuti A, Minnix SL, Carson JL, Hazucha M, Zariwala MA, Knowles MR. Primary ciliary dyskinesia: diagnostic and phenotypic features. *Am J Respir Crit Care Med* 2004;169:459–467.

31 Papon JF, Bassinet L, Cariou-Patron G, Zerah-Lancner F, Vojtek AM, Blanchon S, Crestani B, Amselem S, Coste A, Housset B, *et al*. Quantitative analysis of ciliary beating in primary ciliary dyskinesia: a pilot study. *Orphanet J Rare Dis* 2012;7:78.

32 Piacentini GL, Bodini A, Peroni D, Rigotti E, Pigozzi R, Pradal U, Boner AL. Nasal nitric oxide for early diagnosis of primary ciliary dyskinesia: practical issues in children. *Respir Med* 2008;102:541–547.

33 Pifferi M, Bush A, Maggi F, Michelucci A, Ricci V, Conidi ME, Cangiotti AM, Bodini A, Simi P, Macchia P, *et al*. Nasal nitric oxide and nitric oxide synthase expression in primary ciliary dyskinesia. *Eur Respir J* 2011;37:572–577.

34 Wodehouse T, Kharitonov SA, Mackay IS, Barnes PJ, Wilson R, Cole PJ. Nasal nitric oxide measurements for the screening of primary ciliary dyskinesia. *Eur Respir J* 2003;21:43–47.

35 Santamaria F, De Stefano S, Montella S, Barbarano F, Iacotucci P, Ciccarelli R, Sofia M, Maniscalco M. Nasal nitric oxide assessment in primary ciliary dyskinesia using aspiration, exhalation, and humming. *Med Sci Monit* 2008;14:CR80–CR85.

36 Olin JT, Burns K, Carson JL, Metjian H, Atkinson JJ, Davis SD, Dell SD, Ferkol TW, Milla CE, Olivier KN, *et al*.; Genetic Disorders of Mucociliary Clearance Consortium. Diagnostic yield of nasal scrape biopsies in primary ciliary dyskinesia: a multicenter experience. *Pediatr Pulmonol* 2011;46:483–488.

37 Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW. Primary ciliary dyskinesia: recent advances in diagnostics, genetics, and characterization of clinical disease. *Am J Respir Crit Care Med* 2013;188:913–922.

38 Simoneau T, Zandieh SO, Rao DR, Vo P, Palm KE, McCown M, Kopel LS, Dias A, Casey A, Perez-Atayde AR, *et al*. Impact of cilia ultrastructural examination on the diagnosis of primary ciliary dyskinesia. *Pediatr Dev Pathol* 2013;16:321–326.

39 Boaretto F, Snijders D, Salvoro C, Spalletta A, Mostacciuolo ML, Collura M, Cazzato S, Girosi D, Silvestri M, Rossi GA, *et al*. Diagnosis of primary ciliary dyskinesia by a targeted next-generation sequencing panel: molecular and clinical findings in Italian patients. *J Mol Diagn* 2016;18:912–922.

40 Shapiro AJ, Zariwala MA, Ferkol T, Davis SD, Sagel SD, Dell SD, Rosenfeld M, Olivier KN, Milla C, Daniel SJ, *et al*.; Genetic Disorders of Mucociliary Clearance Consortium. Diagnosis, monitoring, and treatment of primary ciliary dyskinesia: PCD Foundation consensus recommendations based on state of the art review. *Pediatr Pulmonol* 2016;51:115–132.

41 Knowles MR, Ostrowski LE, Leigh MW, Sears PR, Davis SD, Wolf WE, Hazucha MJ, Carson JL, Olivier KN, Sagel SD, *et al*. Mutations in *RSPH1* cause primary ciliary dyskinesia with a unique clinical and ciliary phenotype. *Am J Respir Crit Care Med* 2014;189:707–717.