Canadian Respiratory Journal

WILEY

RESEARCH ARTICLE OPEN ACCESS

Identifying Pediatric Primary Ciliary Dyskinesia in Health Administrative Databases

Wallace B. Wee[1,2,3,4] | Giles Santyr[5,6] | Teresa To[2,3] | Sharon D. Dell[2,3,4]

[1]Division of Respiratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada | [2]Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Ontario, Canada | [3]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada | [4]Division of Respiratory Medicine, Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, Canada | [5]Translational Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada | [6]Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

Correspondence: Wallace B. Wee (wallace.wee@cw.bc.ca)

Received: 23 June 2026 | **Revised:** 3 September 2026 | **Accepted:** 1 November 2026

Academic Editor: Youfeng Zhu

Keywords: case ascertainment algorithm | pediatrics | population-based research | primary ciliary dyskinesia

ABSTRACT

Background: Primary ciliary dyskinesia (PCD) is a rare genetic disorder affecting the cilia. Identifying patients with PCD is challenging due to the heterogeneous, nonspecific, clinical manifestations and the lack of a diagnostic gold standard. A PCD case ascertainment algorithm (PCD-CAA) offers a cost-effective and efficient method for identifying potential PCD patients for further diagnostic testing.

Objective: To develop and assess the performance (i.e., sensitivity, specificity, and Youden's index) of a PCD-CAA.

Methods: We performed a retrospective case–control study using population-based health administrative data housed at ICES in Ontario, Canada, from 1996 to 2020. Inclusion criteria: (1) Have a diagnosis of PCD, cystic fibrosis (CF), or asthma and (2) age < 18 years during the study period. Pediatric PCD patients (cases) were identified from the SickKids PCD Registry and linked with health administrative data. Pediatric patients with CF or asthma (controls) were abstracted from health administrative databases. Numerous PCD-CAA were developed, and their performance was assessed using the CombiROC R package.

Results: We identified 126 cases and 1,262,194 controls (CF: 1307, asthma: 1,260,887). 1024 PCD-CAA were assessed. The most sensitive PCD-CAA was "Pick 2 of 5" of bronchiectasis, chronic bronchitis, neonatal respiratory disorders, rhinosinusitis, and laterality defects (sensitivity: 95%, specificity: 97%, and Youden's index: 0.92).

Discussions: To our knowledge, this is the first study to develop a PCD-CAA using population-based health administrative data. The PCD-CAA demonstrated high sensitivity and specificity, making it a valuable tool for population screening and health policy planning. Further validation studies are necessary to confirm its robustness when applied to other institutional PCD cohorts and provincial health administrative databases.

1 | Introduction

Primary ciliary dyskinesia (PCD) is a rare ciliary disease with increasing global recognition [1–3]. Previous studies have estimated the prevalence to be 0.5 to 1-in-10,000 [3], but a recent genomic study conservatively estimated the PCD global prevalence to be 1-in-7554 [4]. This difference in prevalence estimates is an indicator of the ongoing challenges in identifying and diagnosing potential PCD patients and the limitations of current PCD patient registries [3].

PCD patients often present to their primary care provider with common and nonspecific clinical manifestations, such as cough and nasal congestion (Table 1), making it challenging for

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Copyright © 2026 Wallace B. Wee et al. *Canadian Respiratory Journal* published by John Wiley & Sons Ltd.

TABLE 1 | Select primary ciliary dyskinesia clinical manifestations versus age.

Age	Common clinical manifestations based on age
Fetus	Laterality defects (about 50%)
Infancy	Laterality defects, neonatal respiratory distress, wet cough, nasal congestion, chest infection, atelectasis
Childhood	Laterality defects, chronic wet cough, chronic nasal congestion, otitis media, conductive hearing loss, rhinosinusitis, chest infection, atelectasis, bronchiectasis
Adulthood	Laterality defects, chronic wet cough, chronic nasal congestion, otologic dysfunction, rhinosinusitis, chest infection, atelectasis, bronchiectasis, infertility

clinicians to identify potential PCD patients [1, 3]. Diagnostic predictive tools for PCD, such as the PrImary CiliARy DyskinesiA Rule (PICADAR) [5, 6], have been developed to help clinicians identify PCD patients and improve the positive predictive value (PPV) of diagnostic tests (e.g., nasal nitric oxide [nNO] measurement [7], transmission electron microscopy [TEM] [8], genetic panel testing [9], and high-speed video microscopy analysis [10]). However, these symptom-based diagnostic predictive tools are only effective if utilized by the clinician and only applicable on an individual, case-by-case, basis for patients presenting with symptoms suspicious for PCD. Diagnostic delays are common, with a median age of diagnosis in English and European children reported to be 2.6 and 5.3 years, respectively [11, 12], and lead to delays in establishing appropriate PCD management and, potentially, negatively impact long-term outcomes [13].

Strategies used to identify patients with a specific disorder or disease (such as PCD) are broadly known as case ascertainment methods. These methods can be further divided into three types: active, passive, and hybrid [14, 15]. In active case ascertainment, dedicated staff screen all patients and perform confirmatory diagnostic testing on individuals who meet the screening criteria [15]. In passive case ascertainment, case ascertainment algorithms (CAAs) screen health administrative databases to identify individuals with specific criteria but, unlike active case ascertainment, confirmatory diagnostic testing is not performed [14]. Understandably, active case ascertainment methods are not only the most accurate for identifying patients with a specific disease but are also time- and resource-intensive. On the other hand, passive case ascertainment methods are more efficient to implement and cost-effective but less accurate given the lack of confirmatory diagnostic testing. Hybrid case ascertainment methods combine aspects of both active and passive methods and, if performed rigorously, can increase the accuracy while mitigating the limitations of passive case ascertainment methods [15]. Importantly, the performance (e.g., sensitivity and specificity) of passive and hybrid case ascertainment methods is contingent on the health administrative databases used.

Generally, health administrative databases are repositories of health administrative data, which are claims data for every patient encounter with the medical system, such as medical billings and diagnostic codes associated with a visit to the emergency department [15, 16]. In universal single-payer healthcare systems (like Ontario, Canada), all healthcare data are accessible through linkable health administrative databases [16]. This allows for individual patient records to be linked using unique identifiers and provides the ability to track patients across their lifespan [17]. These attributes have made health administrative databases invaluable resources for health services research and can be leveraged for passive case ascertainment, by way of a CAA [15].

CAAs are structured algorithms that identify individuals in a database who meet specific criteria and can be used for estimating disease prevalence, monitoring disease status, and assessing response to treatment [15]. The main benefits of CAAs are their operationalizability, given that they are inexpensive, utilize preexisting health administrative databases, and do not require additional resources from the established healthcare system workflow. However, developing a CAA is challenging because health administrative databases may have inconsistent disease classification nomenclature, health services code revisions over time, or incomplete/missing data [15]. Furthermore, health administrative databases are not built specifically for CAA [18] and often store raw data (i.e., not analysis ready) and require dedicated expertise in, and resources for, data cleaning and management [19]. Given these considerations, assessing the performance (e.g., sensitivity and specificity) of CAA is critical.

Development of a CAA for PCD would be beneficial, given the ongoing challenges with identifying PCD patients and the lack of a PCD-specific International Classification of Diseases (ICD) code (i.e., confirmed PCD patients are only known by their local healthcare institutions). A PCD-CAA would not only provide the ability to broadly screen the general population to identify which individuals should undergo further confirmatory testing but also help health service researchers and policymakers track and understand the needs of PCD patients in order to implement relevant population-level health policies. In 2025, Burns et al. used machine learning (random forest model) and US insurance claims data to screen for PCD patients, but the algorithms created were only able to achieve a minimum sensitivity of 0.71–0.82 (specificity was not calculated) [20]. Their strategy and results were likely limited by the US insurance claims database's susceptibility to selection or misclassification bias. This is different from Canada, where there is a universal, single-payer healthcare system that captures all patient encounters with the healthcare system across their lifespan and corresponding ICD codes.

Therefore, the purpose of this study is to generate and assess the performance of different PCD-CAAs using health administrative databases from Ontario, Canada.

2 | Methodology

2.1 | Overview

Ontario has a pediatric population of approximately 3 million, and across the province, there are more than 400 hospitals. Ontario is a publicly funded healthcare system, and the Ontario Ministry of Health (MOH) funds most diagnostic tests, procedures, physician services, and hospitals. Therefore, all health administrative data (such as diagnoses, procedures, and billing codes) are reliably collected on every Ontario resident by the

Institute for Clinical Evaluative Sciences (formerly known as the ICES), a nonprofit organization funded by MOH. Data are stored in linkable databases using unique encrypted identifiers for each patient.

2.2 | Study Design

We performed a retrospective case–control study using population health administrative data housed at ICES in Ontario, Canada, from 1996 to 2020 to derive a PCD-CAA. The process schematic for developing the PCD-CAA is shown in Figure 1.

2.3 | Health Administrative Databases

Relevant health administrative database used for this project include the (i) Canadian Institute for Health Information Discharge Abstract Database (CIHI-DAD) for hospital admissions, (ii) National Ambulatory Care Reporting System (NACRS) for emergency department visits, (iii) Same-Day Surgery (SDS) for surgical procedures, and (iv) Ontario Health Insurance Plan (OHIP) for outpatient visits.

2.4 | Cohorts: Cases and Controls

Cases were identified from the Hospital for Sick Children PCD Registry if they were < 18 years old during the study period of 1996–2020, and had a confirmed diagnosis of PCD (i.e., biallelic pathogenic/likely pathogenic variants in one known PCD gene, or a classic ciliary ultrastructural TEM defect) [10, 21, 22]. These cases were then linked to the ICES health administrative databases using their OHIP health card number.

The controls (comparator group) were patients who were < 18 years old during the study period of 1996–2020 and had a diagnosis of cystic fibrosis (CF) or asthma using ICD-9/ICD-10 diagnostic codes (i.e., CF: 277.0/E84, Asthma: 493/J45) coded in the ICES health administrative database.

CF and asthma were chosen as the controls, as they have overlapping clinical manifestations with PCD and can often lead to diagnostic confusion.

TABLE 2 | PCD clinical manifestations.

Clinical manifestations	Corresponding cardinal features of primary ciliary dyskinesia [6]
1 Bronchiectasis	
2 Chronic bronchitis	**Early-onset, year-round wet cough**
3 Pneumonia	
4 Nasal polyps	**Early-onset, year-round nasal congestion**
5 Rhinosinusitis, any of: • Rhinitis • Sinusitis	
6 Otic suppurative disorders, any of: • Perforation of the tympanic membranes • Hearing loss • Myringotomy tubes • Otorrhea	
7 Neonatal respiratory disorders, any of: • Neonatal respiratory distress • Congenital pneumonia • Neonatal aspiration syndrome • Other respiratory conditions originating in the perinatal period	**Unexplained neonatal respiratory distress**
8 Laterality defects, any of: • Congenital cardiac malformations • Dextrocardia • Situs inversus • Congenital malformation of spleen	**Laterality defect**

Note: Corresponding primary ciliary dyskinesia cardinal features [6] are in bold.

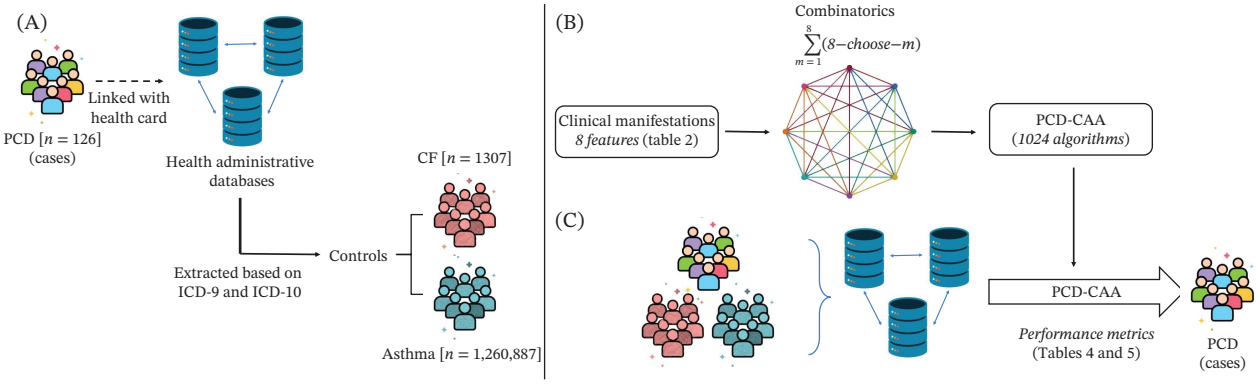


FIGURE 1 | Process schematic for developing the PCD case ascertainment algorithm (PCD-CAA). (A) Linkage of PCD cases to, and identification of controls (i.e., CF and asthma) from, the health administrative databases. (B) PCD-CAA creation using combinatorics of the 8 features, where m ranges from 1 to 8. (C) Application of the PCD-CAA. Abbreviations: CAA = case ascertainment algorithm; CF = cystic fibrosis; PCD = primary ciliary dyskinesia.

TABLE 3 | Health service codes for PCD clinical manifestations.

Clinical manifestation	ICD-10	ICD-9	OHIP	OHIP fee code	CCI	CCP
Bronchiectasis	J47	494	494			
Chronic bronchitis	J40, J41, J42	490, 491	491			
Pneumonia	J12, J13, J14, J15, J16, J17, J18	480, 481, 482, 483, 484, 485, 486, 487.0	486			
Nasal polyps	J33	471	471			
Rhinitis	J30, J31.0	472.0, 477	477			
Sinusitis	J01, J32	461, 473	461, 473			
Perforation of the tympanic membranes	H72	384.2	384			
Hearing loss	H90	389.0, 389.2				
Myringotomy tubes	Z96.22	V43.89		E316, E328, Z912, Z914	1DF53JATS	3201
Otorrhea	H92.1	388.60, 388.69				
Neonatal respiratory distress	P22	769, 770.6, 770.8				
Congenital pneumonia	P23	770.0				
Neonatal aspiration syndrome	P24	770.1				
Other respiratory conditions originating in the perinatal period	P28	770.4, 770.5, 770.7, 770.9				
Congenital cardiac malformations	Q20, Q21, Q22, Q23, Q24.2, Q24.3, Q24.4, Q24.5, Q24.8, Q24.9, Q25, Q26	745, 746, 7470, 7471, 7472, 7473, 7474, 7478, 7479	746			
Dextrocardia	Q24.0	746.87				
Situs inversus	Q89.3	759.3				
Congenital malformation of spleen	Q89.0	759.0	289			

2.5 | Data Extraction

Data from cases and controls were abstracted from the ICES health administrative databases (i.e., CIHI-DAD, NACRS, SDS, and OHIP) and linked using unique encrypted identifiers for each patient. Subsequently, the data were queried for specific clinical manifestations based on the corresponding codes (shown in Tables 2 and 3, respectively). All data extraction was completed using SAS Version 9.4 software (SAS Institute Inc.).

2.6 | PCD-CAA

Clinical manifestations used in the PCD-CAA were selected based on clinical relevance [1] and specificity. Some clinical manifestations were grouped into larger categories (i.e., rhinosinusitis, otic suppurative disorders, neonatal respiratory disorders, and laterality defects) to reduce computational requirements.

PCD-CAA were created using different combinations of clinical manifestations. Each PCD-CAA is defined by three properties: (i) the maximum number of clinical manifestations, (ii) the number of selectable clinical manifestations, and (iii) which clinical manifestations are included. For example, a CAA could have three clinical manifestations (e.g., bronchiectasis, neonatal respiratory disorders, and nasal polyps), of which two are selected. This is denoted as "Pick 2 of 3 (bronchiectasis + neonatal respiratory disorders + nasal polyps)."

2.7 | Statistical Analysis

The performance metrics (sensitivity, specificity, and Youden's index [23]) of each CAA were computed using the R package CombiROC [24–26] and based on case (PCD) versus controls (CF and asthma). Of note, Youden's index was used as a summative metric of the PCD-CAA effectiveness as a diagnostic test because of its versatility in comparing PCD-CAA with differing types and numbers of clinical manifestations. The Youden's index is calculated using the formula "sensitivity + specificity − 1" and ranges from 0 to 1, where 1 is a perfect test and > 0.9 is an excellent test [23]. While the traditional receiver-operator curve (ROC) could be used, it is less appropriate for comparing diagnostic tests that have differing parameters because the ROC is generally used to determine the optimal thresholds for a predefined number of parameters in a diagnostic test.

The top five PCD-CAAs with the highest sensitivities were ranked. The sensitivities between published literature and a PCD-CAA comprised of the same cardinal PCD features [6] were compared. The range of PPVs and negative predictive values (NPVs), also called posttest probabilities, of the selected PCD-CAA were calculated using a range of PCD pretest disease frequency estimates, from 1-in-50 (based on a postulated disease frequency of PCD in symptomatic patients referred for PCD assessment) to 1-in-7554 (global prevalence) [4]. All analyses were performed in R Version 3.6.1 (R Core Team, 2019).

Research ethics board approval was obtained from the Hospital for Sick Children (REB 1000073282).

3 | Results

For the study period of 1996–2020, 126 pediatric PCD patients were found in the Hospital for Sick Children PCD Registry and deterministically linked to ICES using the patients' OHIP numbers. For the same study period, 1307 pediatric CF patients and 1,260,887 pediatric asthma patients were extracted from the ICES databases. A total of 1,262,320 unique patients were included in the study.

The performance metrics of each clinical manifestation are listed in Table 4. The most sensitive were chronic bronchitis (91%) and rhinosinusitis (90%), and the least sensitive were laterality defects (59%) and bronchiectasis (57%). The most specific were bronchiectasis (98%) and nasal polyps (97%), and the least specific was pneumonia (62%).

A total of 1024 PCD-CAA were created, ranging from 1 to 8 clinical manifestations and having different (pick) selections ranging from 1 to the total number of clinical manifestations included. The top five PCD-CAA are shown in Table 5. The PCD-CAA with the highest sensitivity was Pick 2 of 5 (bronchiectasis + chronic bronchitis + neonatal respiratory disorders + rhinosinusitis + laterality defects) with a sensitivity of 95%, specificity of 97%, and Youden's index of 0.92.

TABLE 4 | Performance metrics for each individual clinical manifestation comparing primary ciliary dyskinesia cases versus the comparator cohort (cystic fibrosis and asthma).

Clinical manifestation	Sensitivity (%)	Specificity (%)			Youden's index		
		CF	Asthma	Both	CF	Asthma	Both
Bronchiectasis	57	84	99	98	0.412	0.557	0.556
Chronic bronchitis	91	53	97	97	0.444	0.878	0.878
Nasal polyps	72	63	98	98	0.352	0.707	0.707
Neonatal respiratory disorders*	58	84	93	93	0.423	0.511	0.511
Otic suppurative disorders**	63	89	91	91	0.519	0.537	0.537
Pneumonia	76	37	62	62	0.129	0.379	0.378
Rhinosinusitis***	90	62	97	97	0.516	0.863	0.863
Laterality defects#	62	82	91	91	0.435	0.533	0.532

*Neonatal respiratory disorders as any of neonatal respiratory distress, congenital pneumonia, neonatal aspiration syndrome, or other respiratory conditions originating in the perinatal period.
**Otic suppurative disorders as any of perforation of the tympanic membranes, hearing loss, myringotomy tubes, or otorrhea.
***Rhinosinusitis defined as any of rhinitis or sinusitis.
#Laterality defect defined as any of congenital cardiac malformations, dextrocardia, situs inversus, or congenital malformation of the spleen.

TABLE 5 | Case ascertainment algorithms with the best Youden's index.

Comparator cohort	Pick	Case ascertainment algorithm								Sensitivity (%)	Specificity (%)	Youden's index
		Bronchiectasis	Chronic bronchitis	Nasal polyps	Neonatal respiratory disorders	Otic suppurative disorders	Pneumonia	Rhinosinusitis	Laterality defects			
CF/asthma	**2 of 5**	✓	✓		✓			✓	✓	**95**	**97**	**0.920**
	2 of 5	✓	✓	✓	✓			✓		94	98	0.925
	2 of 4	✓	✓		✓			✓		*94*	*99*	*0.933*
	3 of 7	✓	✓	✓	✓		✓	✓	✓	94	97	0.919
	3 of 6	✓	✓		✓		✓	✓	✓	94	98	0.916

Note: Performance metrics based on primary ciliary dyskinesia cases versus the comparator cohort (cystic fibrosis and asthma). Bold area represents the case ascertainment algorithm with the highest sensitivity. Italic area represents the case ascertainment algorithm with the highest Youden's index.

A PCD-CAA, comprised of PCD criteria–defined cardinal clinical features [6], yielded similar sensitivities to the published literature (Table 6). Specifically, the PCD-CAA of Pick 1-to-4 of 4 (chronic bronchitis + neonatal respiratory disorders + rhinosinusitis + laterality defects) achieved sensitivities of 100%, 93%, 73%, and 35%, respectively.

Using a range of disease frequency estimates (considered as pretest probabilities), from 1-in-50 to 1-in-7554, the calculated posttest probabilities were a PPV ranging from 39.26% to 0.42% and NPV ranging from 99.89% to 100%, when using the most sensitive PCD-CAA (Table 7).

4 | Discussion

To our knowledge, this is the first study to develop a PCD-CAA using population-based health administrative databases from Canada. We found that a PCD-CAA comprised of Pick 2 of 5 (bronchiectasis + chronic bronchitis + neonatal respiratory disorders + rhinosinusitis + laterality defect) had the highest sensitivity. Clinically, this makes sense, as patients with PCD are more likely to experience one, if not all, of these clinical manifestations (i.e., bronchiectasis, chronic bronchitis, neonatal respiratory disorders, rhinosinusitis, and laterality defect) as compared to the general population [1, 6]. In the fetal period, some patients may demonstrate laterality defects on fetal ultrasound. Shortly after birth, many patients will develop unexplained neonatal respiratory distress, often with evidence of rotating atelectasis, requiring supplemental oxygen and hospitalization [27–29]. Before 6 months of age, PCD patients often have chronic bronchitis and chronic nasal congestion [6]. By adulthood, if not earlier, many PCD patients will develop bronchiectasis from recurrent pulmonary infections [30]. Of note, laterality defects are not uncommon in PCD patients [28], and past PCD studies have estimated the prevalence of situs ambiguous and situs inversus totalis in PCD to be ~12–25% and just less than 50%, respectively [31, 32]. Subtle laterality defects may have been missed in our definition, and its performance metrics are likely a conservative estimate. In addition, while patients with bronchiectasis may have chronic bronchitis, it makes sense to include both clinical manifestations because bronchiectasis typically develops later in life and requires chest imaging (which may not be ordered in the primary setting), whereas chronic bronchitis typically presents earlier in life.

It is also not entirely surprising that the other clinical manifestations (i.e., otic suppurative disorders, pneumonia, and nasal polyps) were not included as part of the PCD-CAA. Otic suppurative disorders and pneumonia are not uncommon in the pediatric population [33–35] and are even more prevalent in CF and asthma patients (i.e., our control group). Lastly, while nasal polyps may be a manifestation of an underlying disease (such as CF and PCD), they are most commonly associated with allergic rhinosinusitis [36], and nasal polyps are likely to be underdiagnosed, particularly when the polyps lie deep in the nasal cavity and posterior to the nasal vestibule.

Based on the global prevalence estimate of 1-in-7554, the PPV of the PCD-CAA is 0.42%, and the NPV is 100%. Given that PCD is considered a rare disease, these posttest probabilities are expected. However, the PPV is likely higher as the prevalence estimate of 1-in-7554 is considered conservative due to various

TABLE 6 | Performance metric comparison of case ascertainment algorithm and published cardinal features [6].

Pick	Clinical manifestations	Sensitivity (%)	
		PCD-CAA	Literature [6]
4 of 4	Chronic bronchitis + neonatal respiratory disorders + rhinosinusitis + laterality defects	35	21
3 of 4		73	50
2 of 4		93	80
1 of 4		100	96

study limitations [4]: a limited panel of disease-causing PCD genes (i.e., 29 of > 50 known genes), minimal genetic data for individuals outside of North America and Europe, lack of functional data on the pathogenicity of "variants of unknown significance," and Hardy–Weinberg assumptions. Table 7 gives an example of how the PPV may change with differing pretest probabilities. Moreover, the high NPV of the PCD-CAA suggests that when used as a screening tool, a negative test result would accurately indicate that PCD is unlikely.

There are several strengths of this study. First, the cases were comprised of a well-characterized PCD cohort from the Hospital for Sick Children PCD Registry. Second, the study utilized the entire population of Ontario, Canada, to ascertain the control group (consisting of CF and asthma patients), as opposed to using an institutional control cohort or a small sampling of the province's CF and asthma patients. Third, the study did not require active engagement from healthcare practitioners, eliminating nonresponse bias, which was a significant limitation of earlier clinical epidemiologic studies in PCD [3].

There are limitations to this study. First, the study employed a case–control design with a control cohort that was not a strict population sample, but rather comprised of CF and asthma patients. By selecting controls who are more likely to have clinical manifestations similar to PCD, the specificity of the PCD-CAA is likely underestimated relative to its performance in the general population, and some undiagnosed PCD patients may be captured in the control group. Second, it was assumed that the misclassification of clinical manifestations in the ICES health administrative databases was low, as it is beyond the scope of this project to validate all clinical manifestations. Despite these limitations, we expect that the overall impact on the PCD-CAA would be minimal, as misclassification would only marginally impact the specificity. Finally, the PCD-CAA was developed using ICES databases, which are well-resourced nonprofit organization built for research purposes, and in a highly diverse multicultural population. This may limit the generalizability of the PCD-CAA when applied to different populations, regions, and health administrative databases, as there could be variations in coding standards and diagnostic practices. Therefore, future validation studies in other databases are essential to assess how the PCD-CAA performs in these contexts, and to confirm that its sensitivity and specificity remain consistent before being used in the clinical setting.

Overall, our study is the first to assess the performance of different PCD-CAA using health administrative databases from Canada and determine the optimal combination of clinical manifestations for identifying PCD patients. Our work adds to the literature [20] that PCD-CAA could be a low-cost, efficient way to screen the population for individuals who may benefit from further diagnostic testing. A potential application could be the integration of the PCD-CAA into electronic medical records, and patients who meet the PCD-CAA criteria would be flagged to healthcare providers, prompting timely diagnostic workup and potentially expediting the diagnosis and management of PCD. From a population health perspective, the PCD-CAA has the potential to help policymakers and clinician-researchers better identify PCD patients, assess their burden of disease, and design targeted strategies to enhance care for this population [16]. However, critical next steps include the need for validation studies that assess the robustness and generalizability of the PCD-CAA across diverse patient populations and different healthcare systems. These studies should focus on verifying the algorithm's performance within institutional PCD cohorts and across provincial and national health administrative databases. This is particularly important in resource-limited settings where the infrastructure for health administrative databases and data structures may differ. In addition, exploring the integration of the PCD-CAA into real-time electronic medical records and how it facilitates early intervention could enhance both individual

TABLE 7 | Positive and negative predictive values for the PCD-CAA* (sensitivity 95% and specificity 97%), based on a range of pretest probabilities.

Pre-test probabilities (1-in-*n*)	Positive predictive value (%)	Negative predictive value (%)
50	39.26	99.89
100	24.23	99.95
200	13.73	99.97
500	5.97	99.99
1000	3.07	99.99
5000	0.63	100.00
7554	0.42	100.00

*PCD-CAA Pick 2 of 5 (bronchiectasis + chronic bronchitis + neonatal respiratory disorders + rhinosinusitis + laterality defects).

patient outcomes and broader efforts in public health surveillance.

Nomenclature

CAA	Case ascertainment algorithm
CCI	Canadian Classification of Health Interventions
CCP	Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures
CF	Cystic fibrosis
DAD	Discharge abstract database
ICD	International Classification of Diseases
ICES	Formerly known as the Institute for Clinical Evaluative Sciences
NACRS	National Ambulatory Care Reporting System database
NPV	Negative predictive value
OHIP	Ontario Health Insurance Plan database
PCD	Primary ciliary dyskinesia
PICADAR	PrImary CiliARy DyskinesiA Rule
PPV	Positive predictive value
SDS	Same-Day Surgery database

Acknowledgments

We gratefully acknowledge the support of Emilie Terebessy, Kimball Zhang, and Jingqin Zhu. This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). This document used data adapted from the Statistics Canada Postal CodeOM Conversion File, which is based on data licensed from Canada Post Corporation, and/or data adapted from the Ontario Ministry of Health Postal Code Conversion File, which contains data copied under license from ©Canada Post Corporation and Statistics Canada. Parts of this material are based on data and information compiled and provided by the Ontario Ministry of Health and Ontario Health. The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

Funding

This study was funded by the Royal College of Physicians and Surgeons of Canada.

Disclosure

This study made use of de-identified data from the ICES Data Repository, which is managed by the ICES (formerly known as the Institute for Clinical Evaluative Sciences) with support from its funders and partners: Canada's Strategy for Patient-Oriented Research (SPOR), the Ontario SPOR Support Unit, the Canadian Institutes of Health Research, and the Government of Ontario. The opinions, results and conclusions reported are those of the authors. ICES is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). Parts of this material are based on data and information compiled and provided by the Canadian Institute for Health Information (CIHI). No endorsement by ICES or any of its funders or partners is intended or should be inferred. However, the analyses, conclusions, opinions, and statements expressed herein are those of the author, and not necessarily those of CIHI. Parts of this material are based on data and information provided by Cancer Care Ontario (CCO). The opinions, results, views, and conclusions reported in this paper are those of the authors and do not necessarily reflect those of CCO. No endorsement by CCO is intended or should be inferred. An earlier version of this manuscript was presented as part of Dr. Wee's PhD thesis [37].

Conflicts of Interest

Sharon D. Dell received grant funding from NIH, BCCHRI, and Parion Sciences. Sharon D. Dell has received honoraria from Novartis and AstraZeneca. Sharon D. Dell was a consultant for Sanofi and Regeneron. The remaining authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. W. B. Wee, B. Kinghorn, S. D. Davis, T. W. Ferkol, and A. J. S. Shapiro, "Primary Ciliary Dyskinesia," *Pediatrics* (2024).

2. F. Gahleitner, J. Thompson, C. L. Jackson, et al., "Lower Airway Clinical Outcome Measures for Use in Primary Ciliary Dyskinesia Research: A Scoping Review," *ERJ Open Research* 7, no. 4 (2021): 00320–02021, https://doi.org/10.1183/23120541.00320-2021.

3. W. B. Wee, D. Gatt, E. Seidl, G. Santyr, T. To, and S. D. Dell, "Estimates of Primary Ciliary Dyskinesia Prevalence: A Scoping Review," *ERJ Open Research* 10, no. 4 (2024): 00989–02023, https://doi.org/10.1183/23120541.00989-2023.

4. W. B. Hannah, B. A. Seifert, R. Truty, et al., "The Global Prevalence and Ethnic Heterogeneity of Primary Ciliary Dyskinesia Gene Variants: A Genetic Database Analysis," *Lancet Respiratory Medicine* (2022).

5. L. Behan, B. D. Dimitrov, C. E. Kuehni, et al., "PICADAR: A Diagnostic Predictive Tool for Primary Ciliary Dyskinesia," *European Respiratory Journal* 47, no. 4 (2016): 1103–1112, https://doi.org/10.1183/13993003.01551-2015.

6. M. W. Leigh, T. W. Ferkol, S. D. Davis, et al., "Clinical Features and Associated Likelihood of Primary Ciliary Dyskinesia in Children and Adolescents," *Annals of the American Thoracic Society* 13, no. 8 (2016): 1305–1313, https://doi.org/10.1513/annalsats.201511-748oc.

7. M. W. Leigh, M. J. Hazucha, K. K. Chawla, et al., "Standardizing Nasal Nitric Oxide Measurement as a Test for Primary Ciliary Dyskinesia," *Annals of the American Thoracic Society* 10, no. 6 (2013): 574–581, https://doi.org/10.1513/annalsats.201305-110oc.

8. A. Shoemark, M. Boon, C. Brochhausen, et al., "International Consensus Guideline for Reporting Transmission Electron Microscopy Results in the Diagnosis of Primary Ciliary Dyskinesia (BEAT PCD TEM Criteria)," *European Respiratory Journal* 55, no. 4 (2020): 1900725, https://doi.org/10.1183/13993003.00725-2019.

9. M. A. Zariwala, M. R. Knowles, and M. W. Leigh, "Primary Ciliary Dyskinesia," in *Genereviews(®)*, ed. M. P. Adam (Seattle, WA: University of Washington, 1993).

10. J. S. Lucas, A. Barbato, S. A. Collins, et al., "European Respiratory Society Guidelines for the Diagnosis of Primary Ciliary Dyskinesia," *European Respiratory Journal* 49, no. 1 (2017): 1601090, https://doi.org/10.1183/13993003.01090-2016.

11. C. E. Kuehni, T. Frischer, M. P. Strippoli, et al., "Factors Influencing Age at Diagnosis of Primary Ciliary Dyskinesia in European Children," *European Respiratory Journal* 36, no. 6 (2010): 1248–1258, https://doi.org/10.1183/09031936.00001010.

12. B. Rubbo, S. Best, R. A. Hirst, et al., "Clinical Features and Management of Children With Primary Ciliary Dyskinesia in England," *Archives of Disease in Childhood* 105, no. 8 (2020): 724–729, https://doi.org/10.1136/archdischild-2019-317687.

13. A. Ellerman and H. Bisgaard, "Longitudinal Study of Lung Function in a Cohort of Primary Ciliary Dyskinesia," *European Respiratory Journal*

10, no. 10 (1997): 2376–2379, https://doi.org/10.1183/09031936.97.10102376.

14. A. Mburia-Mwalili and W. Yang, "Birth Defects Surveillance in the United States: Challenges and Implications of International Classification of Diseases, Tenth Revision, Clinical Modification Implementation," *International Scholarly Research Notices* 2014 (2014): 212874–212879, https://doi.org/10.1155/2014/212874.

15. Y. Nava de Escalante, A. Abayomi, S. Langlois, et al., "Validation of Case Definition Algorithms for the Ascertainment of Congenital Anomalies," *Birth Defects Research* 115, no. 3 (2023): 302–317, https://doi.org/10.1002/bdr2.2112.

16. A. S. Gershon, C. Wang, J. Guan, J. Vasilevska-Ristovska, L. Cicutto, and T. To, "Identifying Individuals With Physcian Diagnosed COPD in Health Administrative Databases," *COPD: Journal of Chronic Obstructive Pulmonary Disease* 6, no. 5 (2009): 388–394, https://doi.org/10.1080/15412550903140865.

17. S. E. Andrade, A. Bérard, H. M. Nordeng, M. E. Wood, M. M. van Gelder, and S. Toh, "Administrative Claims Data Versus Augmented Pregnancy Data for the Study of Pharmaceutical Treatments in Pregnancy," *Current Epidemiology Reports* 4, no. 2 (2017): 106–116, https://doi.org/10.1007/s40471-017-0104-1.

18. T. Antoniou, B. Zagorski, M. R. Loutfy, C. Strike, and R. H. Glazier, "Validation of Case-Finding Algorithms Derived From Administrative Data for Identifying Adults Living With Human Immunodeficiency Virus Infection," *PLoS One* 6, no. 6 (2011): e21748, https://doi.org/10.1371/journal.pone.0021748.

19. E. K. Johnson and C. P. Nelson, "Values and Pitfalls of the Use of Administrative Databases for Outcomes Assessment," *Journal d'Urologie* 190, no. 1 (2013): 17–18, https://doi.org/10.1016/j.juro.2013.04.048.

20. G. Burns, C. Kauffman, M. Manion, et al., "Feasibility of Machine Learning Analysis for the Identification of Patients With Possible Primary Ciliary Dyskinesia," *Orphanet Journal of Rare Diseases* 20, no. 1 (2025): 516, https://doi.org/10.1186/s13023-025-03966-z.

21. A. J. Shapiro, S. D. Davis, D. Polineni, et al., "Diagnosis of Primary Ciliary Dyskinesia. An Official American Thoracic Society Clinical Practice Guideline," *American Journal of Respiratory and Critical Care Medicine* 197, no. 12 (2018): e24–e39, https://doi.org/10.1164/rccm.201805-0819st.

22. A. Shoemark, M. Boon, C. Brochhausen, et al., "International Consensus Guideline for Reporting Transmission Electron Microscopy Results in the Diagnosis of Primary Ciliary Dyskinesia (BEAT PCD TEM Criteria)," *European Respiratory Journal* 55, no. 4 (2020): 1900725, https://doi.org/10.1183/13993003.00725-2019.

23. E. F. Schisterman, D. Faraggi, B. Reiser, and J. Hu, "Youden Index and the Optimal Threshold for Markers With Mass at Zero," *Statistics in Medicine* 27, no. 2 (2008): 297–315, https://doi.org/10.1002/sim.2993.

24. S. Mazzara, R. L. Rossi, R. Grifantini, S. Donizetti, S. Abrignani, and M. Bombaci, "Combiroc: An Interactive Web Tool for Selecting Accurate Marker Combinations of Omics Data," *Scientific Reports* 7, no. 1 (2017): 45477, https://doi.org/10.1038/srep45477.

25. M. Bombaci and R. L. Rossi, "Computation and Selection of Optimal Biomarker Combinations by Integrative ROC Analysis Using Combiroc," in *Proteomics for Biomarker Discovery: Methods and Protocols*, ed. V. Brun and Y. Couté (New York, NY: Springer, 2019), 247–259.

26. I. Ferrari, S. Mazzara, M. Crosti, et al., *Combiroc: When 'Less Is More' in Bulk and Single Cell Marker Signatures* (2023).

27. W. B. Wee, M. W. Leigh, and S. D. Davis, "Association of Neonatal Hospital Length of Stay with Lung Function in Primary Ciliary Dyskinesia," *Annals of the American Thoracic Society* (2022).

28. D. A. Fitzgerald and A. J. Shapiro, "When to Suspect Primary Ciliary Dyskinesia in Children," *Paediatric Respiratory Reviews* 18 (2016): 3–7, https://doi.org/10.1016/j.prrv.2015.11.006.

29. T. Mullowney, D. Manson, R. Kim, D. Stephens, V. Shah, and S. Dell, "Primary Ciliary Dyskinesia and Neonatal Respiratory Distress," *Pediatrics* (2014): 2014.

30. P. G. Noone, M. W. Leigh, A. Sannuti, et al., "Primary Ciliary Dyskinesia: Diagnostic and Phenotypic Features," *American Journal of Respiratory and Critical Care Medicine* 169, no. 4 (2004): 459–467, https://doi.org/10.1164/rccm.200303-365OC.

31. A. J. Shapiro, S. D. Davis, T. Ferkol, et al., "Laterality Defects Other Than Situs Inversus Totalis in Primary Ciliary Dyskinesia: Insights Into Situs Ambiguus and Heterotaxy," *Chest* 146, no. 5 (2014): 1176–1186, https://doi.org/10.1378/chest.13-1704.

32. S. Best, A. Shoemark, B. Rubbo, et al., "Risk Factors for Situs Defects and Congenital Heart Disease in Primary Ciliary Dyskinesia," *Thorax* 74, no. 2 (2019): 203–205, https://doi.org/10.1136/thoraxjnl-2018-212104.

33. V. Usonis, T. Jackowska, S. Petraitiene, et al., "Incidence of Acute Otitis Media in Children Below 6 Years of Age Seen in Medical Practices in Five East European Countries," *BMC Pediatrics* 16, no. 1 (2016): 108, https://doi.org/10.1186/s12887-016-0638-2.

34. I. Rudan, K. L. O'Brien, H. Nair, et al., "Epidemiology and Etiology of Childhood Pneumonia in 2010: Estimates of Incidence, Severe Morbidity, Mortality, Underlying Risk Factors and Causative Pathogens for 192 Countries," *Journal of Global Health* 3, no. 1 (2013): 010401, https://doi.org/10.7189/jogh.03.010401.

35. N. Bhattacharyya and S. G. Shay, "Epidemiology of Pediatric Tympanostomy Tube Placement in the United States," *Otolaryngology–Head and Neck Surgery* 163, no. 3 (2020): 600–602, https://doi.org/10.1177/0194599820917397.

36. N. H. Ta, "Will We Ever Cure Nasal Polyps?" *Annals of the Royal College of Surgeons of England* 101, no. 1 (2019): 35–39, https://doi.org/10.1308/rcsann.2018.0149.

37. W. B. Wee, *Primary Ciliary Dyskinesia: An Epidemiological Exploration Furthering Our Understanding of the Burden of Disease in PCD* (ProQuest Dissertations & Theses, 2025).